ICASE

COMPACT FINITE DIFFERENCE SCHEMES FOR MIXED

INITIAL-BOUNDARY VALUE PROBLEMS

Richard B. Philips

and

Milton E. Rose

Report No. 81-18

June 15, 1981

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING

NASA Langley Research Center, Hampton, Virginia 23665

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

COMPACT FINITE DIFFERENCE SCHEMES FOR MIXED INITIAL-BOUNDARY VALUE PROBLEMS

Richard B. Philips*
Naval Underwater Systems Center

and

Milton E. Rose
Institute for Computer Applications in Science and Engineering

ABSTRACT

This paper discusses a class of compact second order accurate finite difference equations for mixed initial-boundary value problems for hyperbolic and convective-diffusion equations. Convergence is proved by means of energy arguments and both types of equations are solved by similar algorithms. For hyperbolic equations an extension of the Lax-Wendroff method is described which incorporates dissipative boundary conditions. Upwind-downwind differencing techniques arise as the formal hyperbolic limit of the convective-diffusion equation. Finally, a finite difference "chain-rule" transforms the schemes from rectangular to quadrilateral subdomains.

This work was performed under NASA Contracts No. NAS1-14472 and NAS1-15810 at ICASE, NASA Langley Research Center, Hampton, VA 23665.

*Former Address: Old Dominion University, Norfolk, VA 23506

INTRODUCTION

This paper discusses a class of implicit compact finite difference schemes which treat mixed initial-boundary value problems in one and two space dimensions for both hyperbolic and convective-diffusion equations. The case in which the equations have constant coefficients serves to illustrate most simply the theoretical results and the algorithmic methods, many of which could be extended with little more than technical complexity to variable coefficients. For example, simple arguments establish energy estimates from which the convergence of the schemes for all fixed values of ratios of time and space mesh parameters immediately follows. Another important result is that the truncation errors for both schemes are second-order in the mesh parameters; in particular, the accuracy of the approximations involved in treating the convective-diffusion equation are independent of the local cell Reynolds number.

Hyperbolic systems of equations arise naturally as the asymptotic "outer expansion" connected with convective-diffusion equations (e.g. Navier-Stokes) and the selection of an appropriate class of weak-solutions of the former to describe shocks, boundary conditions, etc., must have ultimate reference to the latter equations. Two related numerical approaches to hyperbolic equations are described here. One employs dissipative factors $\varepsilon_x, \varepsilon_y$ which modify the nondissipative hyperbolic scheme by an artificial viscosity term. The other treats the hyperbolic problem as the asymptotic limit of the convective-diffusion equation. Each approach appears to allow the hyperbolic problem to be treated in its nonconservative, nonsymmetric form.

The Appendix describes the underlying approximation rationale for the schemes and also explains their relationship to conservative difference forms.

The schemes share a common algorithmic structure. The physics, as generally described by conservation laws, is approximately expressed by a "leap-frog" equation $u_{\cdot}^{n+\frac{1}{2}} = F(u_{\cdot}^{n}, u_{\cdot}^{n-\frac{1}{2}})$ in which $u_{\cdot}^{n} = u(\cdot, t_n)$. The intermediate values $u_{\cdot}^{n}$ are determined by $u_{\cdot}^{n-\frac{1}{2}}$ and the prescribed boundary data at $t = t_n$ by an operator $P: Pu_{\cdot}^{n} = u_{\cdot}^{n-\frac{1}{2}}$. In one-dimension the solution of this algebraic two-point boundary-value problem may be obtained by a method due to H. B. Keller [5], [6]. The treatment of two-dimensional problems can, as will be shown, be approximated by alternating direction methods which employ these one-dimensional techniques.

For hyperbolic problems the solution $u_{\cdot}^{n} = P^{-1}u_{\cdot}^{n-\frac{1}{2}}$ may be approximated to the same accuracy as the difference scheme by the Lax-Wendroff scheme when modified so as to include a completion of boundary conditions in a manner to be described below.

A similar explicit approximation to the solution of the analogous operator equation which arises in the case of the convective-diffusion equation is possible but is not described here. Rather the direct development of a tridiagonal algebraic system which can be used to determine $u_{\cdot}^{n}$ is described in order to help clarify the role of a function $q(\theta)$ of the cell Reynolds number θ which arises in the primary form of the difference equations. The simpler scalar problem illustrates a singular perturbation problem as $|\theta| \to \infty$, the limiting solution of which solves the related reduced hyperbolic problem. For $|\theta| \to \infty$ the function $q(\theta)$ acts as a switch which reduces the above-mentioned tridiagonal system to an upper or a lower bidiagonal system corresponding to an upwind or downwind differencing technique in which one of the parabolic boundary conditions becomes an ineffective constraint for the reduced hyperbolic problem.

The ADI algorithm which yields the solution operator P^{-1} for one-dimensional hyperbolic problems makes essential use of a hyperbolicity assumption which prevents the algorithm from being employed in physical problems in which, say, a change from subsonic to supersonic flow occurs. No such restriction occurs in the corresponding solution algorithm for the convective-diffusion equation, however.

Our principal description of results pertains to rectangular computational cells. In Section 1D a simple "chain-rule" is described which extends these results to quadrilateral cells.

1. HYPERBOLIC PROBLEMS

1A. A Dissipative Compact Scheme

Consider the following strictly hyperbolic mixed initial-boundary value problem for $U = (u_1, u_2, \ldots, u_r)^T$:

$$
\begin{aligned}
&\text{a)} \quad U_t + AU_x + BU_y = 0, && 0 < x, y < 1, \quad 0 < t, \\
&\text{b)} \quad U = U^\circ, && t = 0, \\
&\text{c)} \quad \gamma U = g && \text{on } \Gamma,
\end{aligned}
\tag{1.1}
$$

where Γ is the boundary of the unit square. Unless otherwise stated, A and B are assumed to be symmetric, nonsingular, constant matrices with real eigenvalues and the boundary operator γ to be such that the homogeneous boundary condition $\gamma U = 0$ results in the dissipative boundary conditions $(-1)^{x+1} U^T A U \geq 0$, $x = 0,1$ and $(-1)^{y+1} U^T B U \geq 0$, $y = 0,1$. These conditions allow a simple energy-norm estimate to be given for the solution.

Scheme

We propose the following compact finite difference scheme to treat (1.1) (c.f. Rose [14], Wendroff [16], [17]): with $U_\bullet^n \equiv U_{jk}^n$

$$
\begin{aligned}
&\text{a)} \quad \delta_t U_\bullet^n + A \delta_x U_\bullet^n + B \delta_y U_\bullet^n = 0 \qquad \text{(leap-frog)} \\
&\text{b)} \quad \mu_t U_\bullet^n = \mu_x U_\bullet^n + \varepsilon_x \cdot \frac{\Delta t}{2} \cdot A \, \delta_x U_\bullet^n \\
&\text{c)} \quad \mu_t U_\bullet^n = \mu_y U_\bullet^n + \varepsilon_y \cdot \frac{\Delta t}{2} \cdot B \, \delta_y U_\bullet^n \, .
\end{aligned}
\tag{1.2}
$$

in which ε_x, ε_y are non-negative parameters.

Both the convergence of this scheme and its dissipative character are simple consequences of the following energy argument: multiply (1.2a) by $\mu_t(U_\bullet^n)^T$ and employ (1.2b,c) to obtain

$$0 = \tfrac{1}{2}\Big[\, \delta_t(U_\bullet^n)^T U_\bullet^n + \delta_x(U_\bullet^n)^T A U_\bullet^n + \delta_y(U_\bullet^n)^T B U_\bullet^n \,\Big]$$
$$+ \varepsilon_x \frac{\Delta t}{2} (A\delta_x U_\bullet^n)^T (A\delta_x U_\bullet^n) + \varepsilon_y \frac{\Delta t}{2} (B\delta_y U_\bullet^n)^T (B\delta_y U_\bullet^n) \,.$$

Summing on the spatial indices, employing the fact that $\varepsilon_x, \varepsilon_y$ are non-negative, and recalling that the boundary conditions were assumed to be dissipative there results $\| U_\bullet^n \| \le \| U_\bullet^0 \|$ where $\| U_\bullet^n \|^2 = \Sigma (U_\bullet^n)^T (U_\bullet^n)$ in which the strict inequality applies unless $\varepsilon_x = \varepsilon_y = 0$.

An immediate consequence of this result is: the solutions of the difference scheme (1.2) converge to the solution of (1.1) as the mesh parameters $\Delta x, \Delta y, \Delta t$ tend to zero for any fixed values of the ratios $\lambda_x = \Delta t/\Delta x$, $\lambda_y = \Delta t/\Delta y$. The scheme is nondissipative when $\varepsilon_x = \varepsilon_y = 0$.

Accuracy

The accuracy of the scheme (1.2), as indicated by a truncation error estimate, is most easily examined as follows:

set

$$\tau = \Delta t/2,$$
$$\tau_x^{\pm} = (1 \pm \varepsilon_x)\tau, \qquad \tau_y^{\pm} = (1 \pm \varepsilon_y)\tau, \tag{1.3}$$

and observe that (1.2b) results in

$$\delta_t U_\bullet^n = \tau^{-1}\Big[(\mu_x + \varepsilon_x \tau A \delta_x) U_\bullet^n - U_\bullet^{n-\frac{1}{2}}\Big]$$
$$= \tau^{-1}\Big[U_\bullet^{n+\frac{1}{2}} - (\mu_x + \varepsilon_x \tau A \delta_x) U_\bullet^n\Big] \;; \tag{1.4}$$

similar results obtain from (1.2c). Apply these expressions for $\delta_t U^n$ in (1.2a) to obtain

$$(\mu_x + \tau_x^+ A\delta_x + \tau B\delta_y)U_\bullet^n = U_\bullet^{n-\frac{1}{2}} ,$$

$$(\mu_x - \tau_x^- A\delta_x - \tau B\delta_y)U_\bullet^n = U_\bullet^{n+\frac{1}{2}} ,$$

so that

$$(\mu_x + \tau_x^+ A\delta_x + \tau B\delta_y)U_\bullet^{n+1} - (\mu_x - \tau_x^- A\delta_x - \tau B\delta_y)U_\bullet^n = 0. \tag{1.5}$$

or, more compactly,

$$\delta_t\mu_x U_\bullet^{n+\frac{1}{2}} + \left(A\delta_x + B\delta_y\right)\mu_t U_\bullet^{n+\frac{1}{2}} + \varepsilon_x\tau A\delta_x\delta_t U_\bullet^{n+\frac{1}{2}} = 0.$$

Similarly,

$$\delta_t\mu_y U_\bullet^{n+\frac{1}{2}} + \left(A\delta_x + B\delta_y\right)\mu_t U_\bullet^{n+\frac{1}{2}} + \varepsilon_y\tau B\delta_y\delta_t U_\bullet^{n+\frac{1}{2}} = 0.$$

Clearly,

$$\text{truncation error} = -\varepsilon_x\tau AU_{xt} + O(\tau^2),$$

$$= -\varepsilon_y\tau BU_{yt} + O(\tau^2),$$

and it is evident that the scheme is second order accurate when the dissipative terms $\varepsilon_x, \varepsilon_y$ vanish.

In one dimension $B = 0$ and $U_t = -AU_x$ so that the truncation error is $\varepsilon_x\tau A^2U_{xx}$ which shows a relationship to an artificial viscosity term. The dissipative energy estimate given earlier, however, shows directly that the terms $\varepsilon_x\tau AU_{xt}$ and $\varepsilon_y\tau BU_{yt}$ yield dissipation.

Amplification and Phase Error

Consider the scalar equation $u_t + au_x = 0$ where $a > 0$. Initial data given by $u^0 = \exp(i\theta x)$ are transformed by (1.2) into $v = \rho \exp[i(\theta x - \psi)]$ while the differential equation carries u^0 into $v' = \exp[i\theta(x - a\Delta t)]$. Eq. (1.5) shows that

$$\rho^2 = \frac{1 + (a\lambda^-)^2}{1 + (a\lambda^+)^2} < 1, \tag{1.6}$$

where $\lambda^{\pm} = (1 \pm \varepsilon_x)\lambda$, $\lambda = \Delta t/\Delta x$, while

$$\tfrac{1}{2}\psi = \arctan(a\lambda \tan(\theta/2)). \tag{1.7}$$

Thus $|\rho| < 1$ for $\varepsilon_x > 0$, i.e. (1.2) is dissipative for $\varepsilon_x > 0$.

For $a\lambda = 1$, $\psi = \theta$. Since $\tan \alpha\phi > \alpha \tan \phi$, $0 \leq \alpha < 1$, then $\psi < \theta$ for $0 < a\lambda < 1$; similarly $\psi > \theta$ for $a\lambda > 1$. Thus the wave speed associated with the difference equation is greater or less than the wave speed of the differential equation according as the CFL number is greater or less than unity.

In the non-scalar case an analysis of the amplification matrix associated with (1.2) yields an inequality similar to (1.6). This observation also implies the convergence of the scheme (1.2) since it is, clearly, consistent with the differential equation (1.1).

1B. Solution Methods

We shall find it convenient to introduce separate notations for the values of $U^n_\bullet$ associated with the sides of a computational cell $\pi^n_\bullet$ and thus rewrite (1.2) as

$$\begin{aligned} &\delta_t U^n_\bullet + A\delta_x V^n_\bullet + B\delta_y W^n_\bullet = 0, \\ &\mu_t U^n = \mu_x V^n_\bullet + \varepsilon_x \tau A \delta_x V^n_\bullet \quad , \\ &\mu_t U^n_\bullet = \mu_y W^n_\bullet + \varepsilon_y \tau B \delta_y W^n_\bullet \quad . \end{aligned} \tag{1.2}'$$

Write

$$P_x \overset{\text{def}}{\equiv} (\mu_x + \tau_x^+ A\delta_x),$$

$$P_y \overset{\text{def}}{\equiv} (\mu_y + \tau_y^+ B\delta_y), \tag{1.8}$$

where τ_x^+, τ_y^+ are given by (1.3).

The argument which led to (1.5) allows us to re-express (1.2) as

$$\text{a)} \quad \delta_t U_\bullet^n + A\delta_x V_\bullet^n + B\delta_y W_\bullet^n = 0 ,$$

$$\text{b)} \quad P_x V_\bullet^n + \tau B\delta_y W_\bullet^n = U_\bullet^{n-\frac{1}{2}} , \tag{1.9}$$

$$\text{c)} \quad P_y W_\bullet^n + \tau A\delta_x V_\bullet^n = U_\bullet^{n-\frac{1}{2}} .$$

These equations lead to the following:

Two-step Algorithm

a) with $U_\bullet^{n-\frac{1}{2}}$ given as "initial" conditions, solve (9b,c) for $V_\bullet^n$, $W_\bullet^n$ with the boundary conditions specified by (1.1c), then

b) solve the leap-frog equations (1.9a) for $U_\bullet^{n+\frac{1}{2}}$ and repeat the procedure for the next time step.

Note that dissipation is effected by simply increasing the value of τ in step (a) as (1.3) indicates.

A formal solution to (9b,c) which is accurate to $0(\tau^2)$, the same as the order of accuracy of the difference equations (1.9), is given by

$$V_\bullet^n = P_x^{-1}(I - \tau B\delta_y P_y^{-1})U_\bullet^{n-\frac{1}{2}}$$

$$W_\bullet^n = P_y^{-1}(I - \tau A\delta_x P_x^{-1})U_\bullet^{n-\frac{1}{2}} \tag{1.10}$$

The following discussion will clarify that the fact that P_x^{-1}, P_y^{-1} may each be obtained as the solution of one-dimensional two-point algebraic boundary value problems.

1C. The Solution Operator P^{-1}

Consider the one-dimensional problem arising from $B = 0$ in (1.1). The difference equations (1.9) then reduce to

$$\text{(1.11)} \qquad \begin{aligned} &\text{a)} \quad \delta_t U^n + A\delta_x V^n_\bullet = 0, \\ &\text{b)} \quad P_x V^n_\bullet = U^{n-\frac{1}{2}}_\bullet, \end{aligned}$$

in which the boundary conditions (1.1c) are incorporated as, say, $\gamma_A^+ U(0,y,t) = g(0,t)$ and $\gamma_A^- U(1,y,t) = g(1,t)$ in which, if A has k_A^+ eigenvalues and k_A^- negative eigenvalues, then rank $\gamma_A^+ = k_A^+$, rank $\gamma_A^- = k_A^-$. Equation (1.11b) together with these boundary conditions may be written in the form

$$\text{(1.12)} \qquad \begin{aligned} \gamma^+ U_0^n &= g(0,t^n), \\ \tfrac{1}{2}(I - \lambda_x A)U^n_{i-\frac{1}{2}} + \tfrac{1}{2}(I + \lambda_x A)U^n_{i+\frac{1}{2}} &= U_i^{n-\frac{1}{2}}, \quad i=\tfrac{1}{2},\tfrac{3}{2},\ldots,M-\tfrac{1}{2}, \\ \gamma^- U_M^n &= g(1,t^n), \end{aligned}$$

the solution of which, as an algebraic two-point boundary value problem, has been described by H. B. Keller [5], [6] in terms of a block-tridiagonal systems of equations (c.f. deBoor and Weiss [2], Lentini and Pereyra [9]).

An explicit approximation to P_x^{-1} may also be given in terms of the Lax-Wendroff method which, if

$$\text{(1.13)} \qquad Q_x \overset{\text{def}}{=} (\mu_x - \tau A\delta_x),$$

is given by

$$\delta_t U^n + A\delta_x V^n_\bullet = 0,$$

(1.14)

$$V^n_\bullet = Q_x U^{n-\frac{1}{2}}_\bullet$$

for the initial value problem for $U_t + AU_x = 0$. We recall that the Lax-Wendroff scheme is dissipative ([11]).

Consider (1.11) in the nondissipative case $\varepsilon_x = 0$, so that

$$P_x = (\mu_x + \tau A\delta_x), \tag{1.15}$$

and observe that, formally,

$$Q_x P_x = I + \frac{\Delta x^2}{4}(I - \lambda_x^2 A^2)\delta_x^2, \tag{1.16}$$

where $\lambda_x = \Delta t/\Delta x$. Multiplying (1.11b) by Q_x and incorporating the boundary conditions (1.1c) there results the equivalent system

$$\text{a)} \quad \begin{cases} \gamma_A^+ V_0^n = g(0,t_n) \ , \\ \gamma_A^- P_x V^n_{\frac{1}{2}} = \gamma_A^- U^{n-\frac{1}{2}}_{\frac{1}{2}} \ , \end{cases}$$

$$\text{b)} \quad Q_x P_x V^n = Q_x U^{n-\frac{1}{2}}_\bullet \ , \tag{1.17}$$

$$\text{c)} \quad \begin{cases} \gamma_A^+ P_x V^n_{M-\frac{1}{2}} = \gamma_A^+ U^{n-\frac{1}{2}}_{M-\frac{1}{2}} \ , \\ \gamma^- V_M^n = g(1,t_n) \ . \end{cases}$$

Equations (1.17a,c) may be solved for V_0^n and V_M^n with the result, say,

$$
\begin{aligned}
&\text{a)} \qquad V_0^n = G_0(V_{1,}^n U_{\frac{1}{2}}^{n-\frac{1}{2}}) \quad , \\
&\text{b)} \qquad V_M^n = G_1(V_{M-1,}^n U_{M-\frac{1}{2}}^{n-\frac{1}{2}}).
\end{aligned} \tag{1.18}
$$

Next, as is most easily verified by considering the case in which A is expressed in diagonal form, for $|\lambda_x A| \leq 1$ the solution of (1.17b) may be approximated with second order accuracy by

$$
V_k^n = Q_x U_k^{n-\frac{1}{2}} \qquad , \qquad k=1,2,\ldots,M-1, \tag{1.19}
$$

following which the boundary values V_0^n, V_M^n can be obtained from (1.18).

Let $\hat{Q}_x$ indicate the operator Q_x extended so as to include the boundary values determined by (1.18). Then the extended Lax-Wendroff approximation to P_x^{-1} is described by

$$
\begin{aligned}
\delta_t U_\cdot^n + A\delta_x V_\cdot^n &= 0, \\
V_\cdot^n &= \hat{Q}_x U_\cdot^{n-\frac{1}{2}} .
\end{aligned} \tag{1.20}
$$

Recalling earlier remarks we conclude: the solution of (1.11) in the nondissipative case can be approximated to $O(\tau^2)$ by the extended Lax-Wendroff operator $\hat{Q}_x$ if $\lambda_x = \Delta t/\Delta x$ is restricted by the CFL condition $|\lambda_x A| \leq 1$; however, the resulting approximation to (1.11) will be dissipative.

By employing $P_x^{-1} \simeq \hat{Q}_x$, $P_y^{-1} \simeq \hat{Q}_y$ in (1.10), combining with (1.9a), and omitting terms of $O(\tau^3)$ the two-dimensional form of the Lax-Wendroff scheme results (now generalized to incorporate boundary conditions).

1D. Difference Equations on Quadrilateral Cells - a Transformation Rule

A simple finite difference analogue of the differential chain rule can be given which allows the compact difference scheme which has been described for rectangular computational cells to be applied on quadrilateral or triangular cells. The transformation does not affect the accuracy of the difference scheme.

Assume that the underlying domain D is partitioned into quadrilateral subdomains; a typical such cell $\hat{\pi}$ whose vertices are the points Q_1, Q_2, Q_3 and Q_4 is indicated in Figure 1.

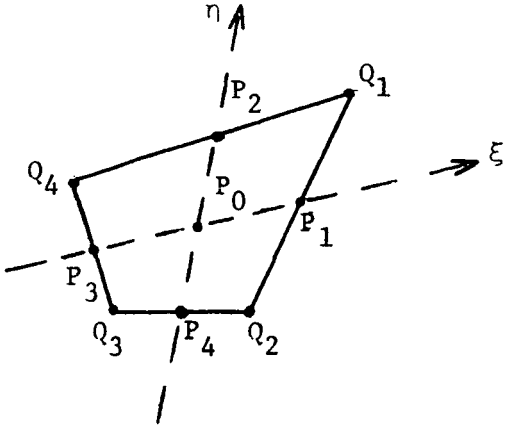

Figure 1. Local cell coordinates (ξ,η) in the cell $\hat{\pi}$.

The area of $\hat{\pi}$ is $\hat{\omega} \equiv \omega(\hat{\pi}) = \frac{1}{2}(\overline{Q_1Q_3} \times \overline{Q_4Q_2})$. As indicated in the figure the points P_1, P_2, P_3, and P_4 are the midpoints of the sides of $\hat{\pi}$ and P_0 is the point of intersection of P_1P_3 and P_2P_4.

The (x,y) coordinates of the points P_i will be denoted by (x_i,y_i), $i=0,1,2,3,4$. With P_0 as center introduce new coordinates $\xi = \xi(x,y)$, $\eta = \eta(x,y)$ with axes as indicated in Figure 1 and with respect to which we may also write

$$P_i = P(\xi_i,\eta_i), \qquad i=0,1,\ldots,4, \quad \text{with} \quad \xi_0 = \eta_0 = 0.$$

We employ the following notations for central difference and averaging operators:

$$\begin{aligned} \Delta_\xi W_0 &= (W(P_1) - W(P_3)), \\ \mu_\xi W_0 &= (W(P_1) + W(P_3))/2, \\ \Delta_\eta W_0 &= (W(P_2) - W(P_4)), \\ \mu_\eta W_0 &= (W(P_2) + W(P_4))/2 . \end{aligned} \tag{1.21}$$

We also set

$$J \overset{\text{def}}{\equiv} \begin{pmatrix} \Delta_\eta y & -\Delta_\xi y \\ -\Delta_\eta x & \Delta_\xi x \end{pmatrix} \div \omega \quad , \tag{1.22}$$

where

$$\omega \overset{\text{def}}{\equiv} (\Delta_\xi x \cdot \Delta_\eta y - \Delta_\eta x \cdot \Delta_\xi y). \tag{1.23}$$

Thus

$$J^{-1} = \begin{pmatrix} \Delta_\xi x & \Delta_\xi y \\ \Delta_\eta x & \Delta_\eta y \end{pmatrix} \quad . \tag{1.24}$$

We shall now show that, when $\varepsilon_x = \varepsilon_y = 0$, the difference equations (1.2) which provides a second-order accurate approximation to (1.1) for a rectangular grid may be expressed in terms of values at the points $P_0, P_1, \ldots, P_4$ of the quadrilateral by

$$\text{a)} \quad \left[\delta_t + (A,B)J\begin{pmatrix}\Delta_\xi \\ \Delta_\eta\end{pmatrix}\right] U(P_0,t_n) = 0,$$

$$(\mu_t - \mu_\xi)U(P_0,t_n) = 0, \tag{1.25}$$

$$\text{b)} \quad (\mu_t - \mu_\eta)U(P_0,t_n) = 0.$$

Comparing (1.25a) with (1.2) we see that the transformation between the rectangular and quadrilaterals cells is given by the "chain rule"

$$\begin{pmatrix}\delta_x \\ \delta_y\end{pmatrix} = J\begin{pmatrix}\Delta_\xi \\ \Delta_\eta\end{pmatrix} \quad . \tag{1.26}$$

The verification of this result depends upon the same argument as that which leads to (1.2) in the case of rectangular cells and which is described in the Appendix. We first observe that a simple class of solutions of the differential equation (1.1) in the quadrilateral cell $\hat{\pi}$ for $|t - t_n| < \Delta t/2$ is given by

$$U(x - x_0,\ y - y_0\ t - t_n) = \underline{a} + [(x - x_0)I - (t - t_n)A]\underline{b} + [(y - y_0)I - (t - t_n)B]\underline{c} \ ,$$

where $\underline{a}$, $\underline{b}$, $\underline{c}$ are arbitrary vector parameters associated with the approximation in $\hat{\pi}$. The points (P_i,t_n), $i = 1, 2, 3, 4$ and the points $(P_0, t_{n-\frac{1}{2}})$

and $(P_0, t_{n+\frac{1}{2}})$ are center points of the faces of the cylinder set $[\pi\chi(|t - t_n| < \Delta t/2)]$. Hence

$$U(P_i,t_n) = a + (x_i - x_0)\underline{b} + (y_i - y_0)\underline{c} \qquad i=1,2,3,4.$$

$$U(P_0, t_{n\pm\frac{1}{2}}) = \underline{a} \mp \frac{\Delta t}{2}[A\underline{b} + B\underline{c}].$$

The result of eliminating the parameters $\underline{a}$, $\underline{b}$, $\underline{c}$ from these expressions is just (1.25).

Equations (1.25) may also be expressed (by the argument which led to (1.5)) by

$$\left[\mu_\xi\delta_t + (A,B)J \cdot \begin{pmatrix}\Delta_\xi \\ \Delta_\eta\end{pmatrix}\right]\mu_t U(P_0,t_n) = 0,$$

$$\left[\mu_\eta\delta_t + (A,B)J \cdot \begin{pmatrix}\Delta_\xi \\ \Delta_\eta\end{pmatrix}\right]\mu_t U(P_0,t_n) = 0.$$

Observing that

$$J \cdot \begin{pmatrix}\Delta_\xi \\ \Delta_\eta\end{pmatrix} = \begin{pmatrix}\partial_x \\ \partial_y\end{pmatrix} + 0(|\hat{\omega}|),$$

the second order accuracy of the approximation on $\hat{\pi}$ is immediately apparent.

1E. Numerical Experiments

This section reports on a number of numerical experiments which were conducted to validate the theory described above. The numerical results given for certain Riemann problems indicates that the scheme (1.2) may be extended to nonlinear problems in which A and B need not be symmetric. All of the methods reported upon employed the construction of P^{-1} indicated by (1.12).

Experiment 1.1

The equation $U_t + AU_x + BU_y = 0$ with $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ was solved using (1.2) with the initial condition

$$U(x,y,0) = \begin{pmatrix} \cos x + \cos y \\ \cos x + \cos y \end{pmatrix},$$

with boundary conditions

$$u_2(0,y,t) = \gamma_A^+ U(0,y,t) = \cos t + \cos(y+t),$$

$$u_1(1,y,t) = \gamma_A^- U(1,g,t) = \cos(1+t) + \cos(y+t),$$

$$u_1(x,0,t) = \gamma_B^+ U(x,0,t) = \cos(x+t) + \cos(t),$$

$$u_2(x,1,t) = \gamma_B^- U(x,1,t) = \cos(x-t) + \cos(1+t).$$

The analytic solution is

$$U(x,y,t) = \begin{pmatrix} \cos(x+t) + \cos(y+t) \\ \cos(x-t) + \cos(y+t) \end{pmatrix}.$$

Table I illustrates the maximum value of the errors for the components of $U = (u_1, u_2)$. These results show the quadratic convergence expected. Mitchell (c.f. [10], Table 10, p. 189) describes experimental results for this equation employing an ADI method and the Lax-Wendroff method in which both components of the solution were specified at the boundary. His results appear to be considerably more accurate than those reported here, a fact, we believe, which is due to his employing boundary data given by the analytical solution. This example emphasizes the important influence the treatment of boundary conditions can have upon solution methods.

Table I

	parameter values	number of time steps	maximum errors u_1	u_2
a)	$\Delta x = \Delta y = 0.1$	10	0.386×10^{-1}	0.377×10^{-1}
	$\Delta t = 0.1$	20	0.233×10^{-1}	0.226×10^{-1}
		40	0.378×10^{-1}	0.370×10^{-1}
b)	$\Delta x = \Delta y = 0.05$	10	0.160×10^{-1}	0.160×10^{-1}
	$\Delta t = 0.05$	20	0.190×10^{-1}	0.188×10^{-1}
		40	0.114×10^{-1}	0.113×10^{-1}

Experiment 1.2

Equations (1.9) are also equivalent to the system of equations

$$\text{a)} \quad P_x V_\bullet^n = U_\bullet^{n-\frac{1}{2}} - \tau B \delta_y W_\bullet^n ,$$

(1.28)

$$\text{b)} \quad \mu_x V^n = \mu_y W_\bullet^n ,$$

$$\text{c)} \quad \delta_t U^n + A\delta_x V_\bullet^n + B\delta_y W_\bullet^n = 0.$$

Employing (1.28b), then

$$\delta_y W_{ij}^n = \frac{2}{\Delta y} (\mu_x V_{ij}^n - W_{i,j-\frac{1}{2}}^n), \tag{1.29}$$

so that (1.28a) is equivalent, when $\varepsilon_x = \varepsilon_y = 0$, to

$$\left[\left(1 + (\frac{\Delta t}{\Delta y})B\right)\mu_x + \tau A \delta_x\right] V_{ij}^n = U_{ij}^{n-\frac{1}{2}} + \frac{\Delta t}{\Delta y} B \, W_{i,j-\frac{1}{2}}^n . \tag{1.30}$$

For scalar equations (1.28b, c) and (1.30) permit the solution to be explicitly determined since, if $W_{\bullet,j-\frac{1}{2}}^n$ is known, $V_{\bullet,j}^n$ may be determined from (1.30) and $W_{\bullet,j+\frac{1}{2}}^n$ from (1.28b), from which $U_\bullet^{n+\frac{1}{2}}$ is then determined from (1.28c).

Table II describes numerical results employing this scheme for the scalar equation $u_t + u_x + 2u_y = 0$ with the initial and boundary conditions

$$u(x,y,0) = \cos x + \cos y \quad , \quad t = 0$$

$$u(0,y,t) = \cos t + \cos(y - 2t),$$

$$u(x,0,t) = \cos(x - t) + \cos 2t ,$$

the analytic solution of which is $u = \cos(x - t) + \cos(y - 2t)$.

Table II

	parameter values	number of time steps	maximum error
a)	$\Delta x = \Delta y = 0.1$	10	0.7×10^{-2}
	$\Delta t = 0.1$	20	0.545×10^{-2}
		40	0.465×10^{-2}
b)	$\Delta x = \Delta y = 0.05$	10	0.215×10^{-2}
	$\Delta t = 0.05$	20	0.173×10^{-2}
		40	0.135×10^{-2}
c)	$\Delta x = \Delta y = 0.1$	10	0.301×10^{-1}
	$\Delta t = 0.2$	20	0.229×10^{-1}
		40	0.205×10^{-1}
d)	$\Delta x = \Delta y = 0.05$	10	0.766×10^{-2}
	$\Delta t = 0.1$	20	0.709×10^{-2}
		40	0.514×10^{-2}

Experiment 1.3

Under suitable boundary conditions (1.1) may have a steady-state solution which may be calculated from (1.2) by letting $n \to \infty$. According to (1.2) the difference equations

$$\text{(1.31)} \qquad \begin{aligned} &\text{a)} \quad A\delta_x V_{ij} + B\delta_y W_{ij} = 0, \\ &\text{b)} \quad \mu_j W_{ij} = \mu_x V_{ij}, \end{aligned}$$

may be expected to describe the steady-state solution more directly. Employing (1.29) in (1.31a) we see that (1.31) is equivalent to the system

$$\text{(1.31)'} \qquad \begin{aligned} (B\mu_x + \frac{\Delta y}{2} A\delta_x)V_{ij} &= BW_{i,j-\frac{1}{2}}, \\ \mu_x V_{ij} &= \mu_x W_{ij}. \end{aligned}$$

For scalar problems, these equations yield the steady-state solution in an explicit manner.

Table III illustrates the results of employing (1.31)' to treat the scalar equation $u_t + au_x + bu_y = 0$ with the boundary conditions $u(0,y,t) = \exp(-y/b)$, $u(x,0,t) = \exp(x/a)$ which are assoicated with the steady-state solution $u(x,y) = \exp(x/a - y/b)$.

Table III

parameter values	maximum error
a=b=1; $\Delta x = \Delta y = 0.1$	0.307×10^{-2}
$\Delta x = \Delta y = 0.05$	0.808×10^{-3}
a=1,b=2; $\Delta x = \Delta y = 0.1$	0.123×10^{-2}
$\Delta x = \Delta y = 0.05$	0.332×10^{-3}

The last two examples illustrate numerical treatments which are possible for the scalar equation but which cannot be directly applied to systems of equations. However, they suggest that a suitable extention of these methods may also be adapted to treat systems of equations.

Experiment 1.4: Riemann Problems

For nonlinear one-dimensional problems, $A = A(U)$ and it is natural to apply (1.2) in which the coefficient A is determined by $\mu_x A(U^n_\cdot)$ or $A(\mu_x U^n_\cdot)$. Because (1.2) is equivalent to an artificial viscosity method it may be expected that the dissipative scheme will converge to the physical weak solution of the nonlinear conservation system $U_t + F_x(U) = 0$ ([11]). The relationship of (1.2) to this conservation equation is described in the Appendix.

In one-dimension the nonconservation form of the Euler equations for inviscid fluid flow is described by

$$U_t + AU_x = 0,$$

where $U = (\rho,u,p)'$ (ρ = density, u = velocity, p = pressure) and

(1.32)
$$A = \begin{pmatrix} u & \rho & 0 \\ 0 & u & \rho^{-1} \\ 0 & \gamma p & u \end{pmatrix},$$

($\gamma = 1.4$).

The results of several numerical experiments with Riemann problems employing the two-step method will be presented here.

Figure 2 illustrates the numerical density profile of a shock traveling to the right with speed 0.979 which results from the initial conditions

$x < 0$	$x > 0$
$\rho = 0.313$	0.219
$u = 0.3$	0.0
$p = 0.166$	0.1

The indicated values of ρ and u on the left and p on the right were used to supply boundary condtions. The dissipation factor in (1.2) was $\varepsilon = 0.15\ \lambda^{-1}$ and $\Delta x = .01$. The value $\lambda = 1.04$ in Figure 2a approximates the situation in which the average value of the CFL numbers before and after the shock was 1. The smoothness of the transition across the shock and the fairly accurate tracking of the correct shock position (indicated by the vertical line) are evident. Figure 2b illustrates the result of increasing the CFL number on both sides of the shock ($\lambda = 1.3$) while in Figure 2c the average CFL number was reduced ($\lambda = 0.7$). The post shock oscillation when $\lambda = 1.3$ and the preshock oscillation when $\lambda = 0.7$ which was evident in the figures may be interpreted as the influence of the CFL number on the wave velocity as discussed in 1A.

Figure 3 illustrates the ρ, u, p profiles resulting from the initial conditions

$x < 0$	$x > 0$
$\rho = 1.0$	0.125
$u = 0.0$	0.0
$p = 1.0$	0.125

The analytical solution, represented in the figures by the continuous line, yields a shock with speed 1.822 and a contact with speed 0.878. The calculated values at time = 2.4 with $\Delta x = 0.1$, $\lambda = 0.6$, and a dissipation factor of $\varepsilon = .125\lambda^{-1}$ yielded fairly good agreement with the exact solution. The relative error of the shock speed was estimated to be 3% and the relative

error of the total energy was calculated to be 0.3%. The value $\lambda = 0.6$ chosen results in an average CFL number $\simeq 1$ through the shock zone. In this experiment the matrix A in (1.2) was estimated by $A_i^n = \mu_x A(U_i^n)$.

In both of the above experiments the dissapative factor ε was kept constant for all spatial points. (As noted by a reviewer, it would have been preferable to have chosen ε to be independent of λ.)

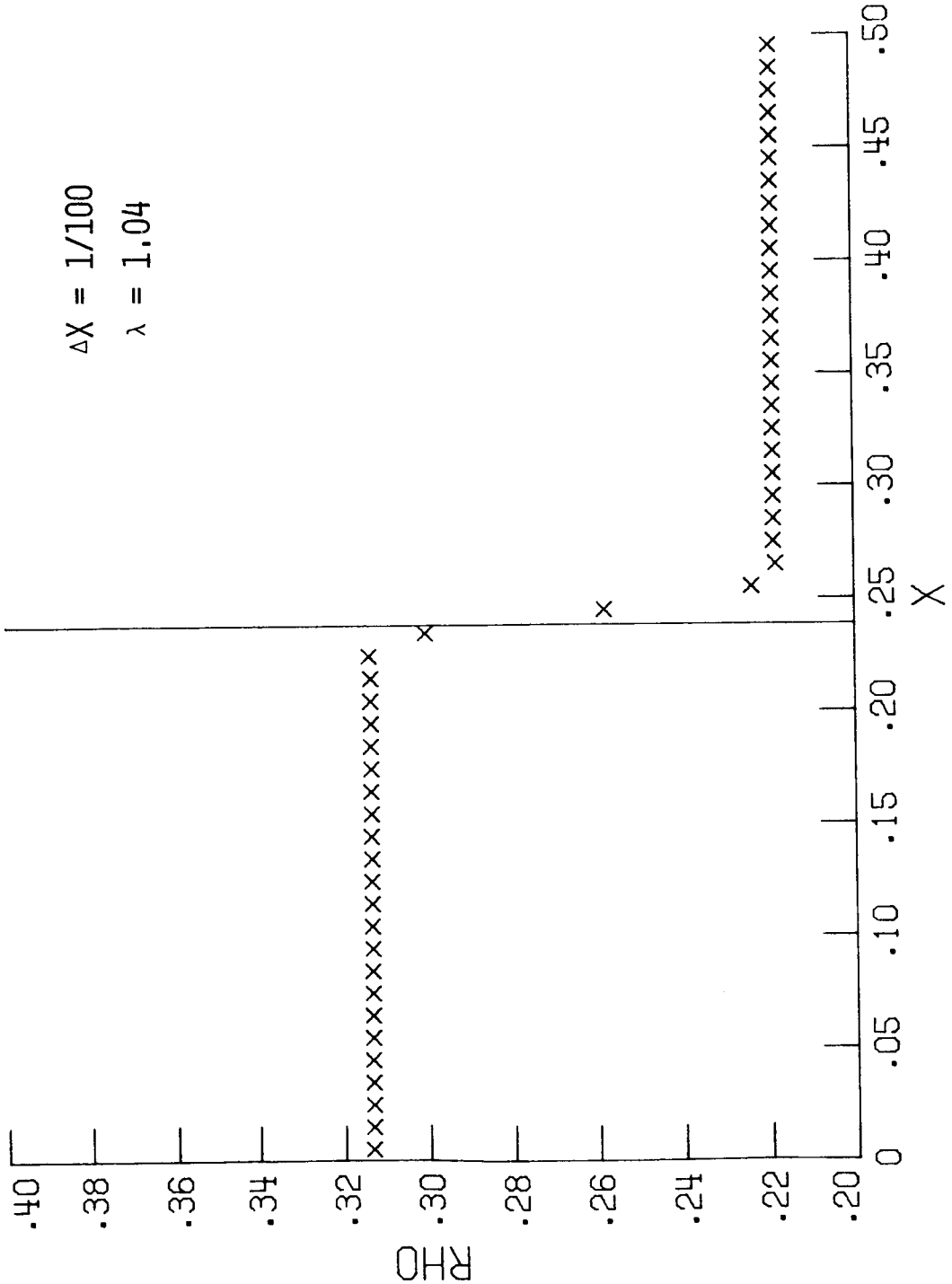

Figure 2a. Density profile of a shock calculation showing the influence of the CFL number upon pre- and post- shock oscillations (Scheme 1.2).

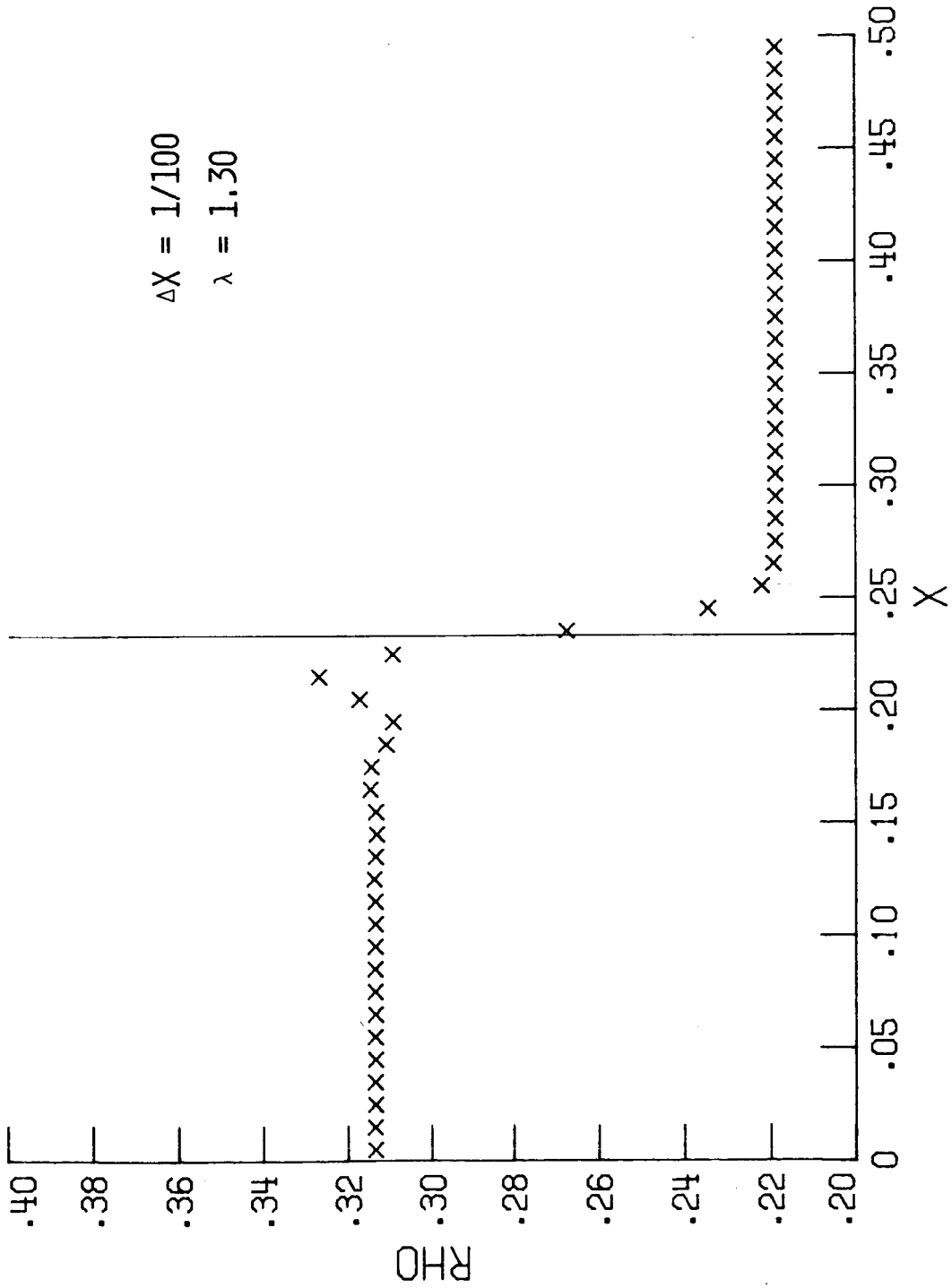

Figure 2b. Density profile of a shock calculation showing the influence of the CFL number upon pre- and post- shock oscillations (Scheme 1.2).

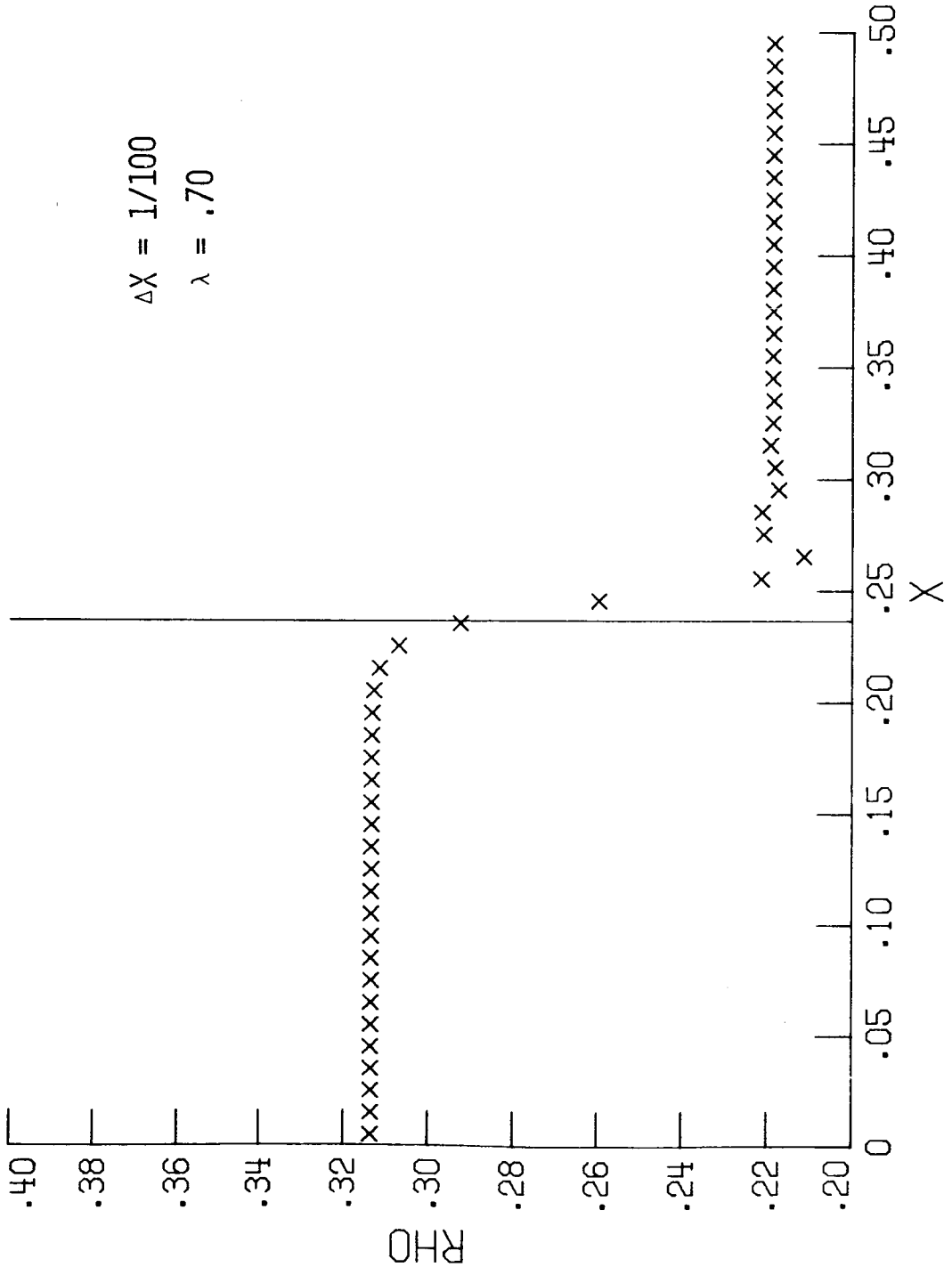

Figure 2c. Density profile of a shock calculation showing the influence of the CFL number upon pre- and post- shock oscillations (Scheme 1.2).

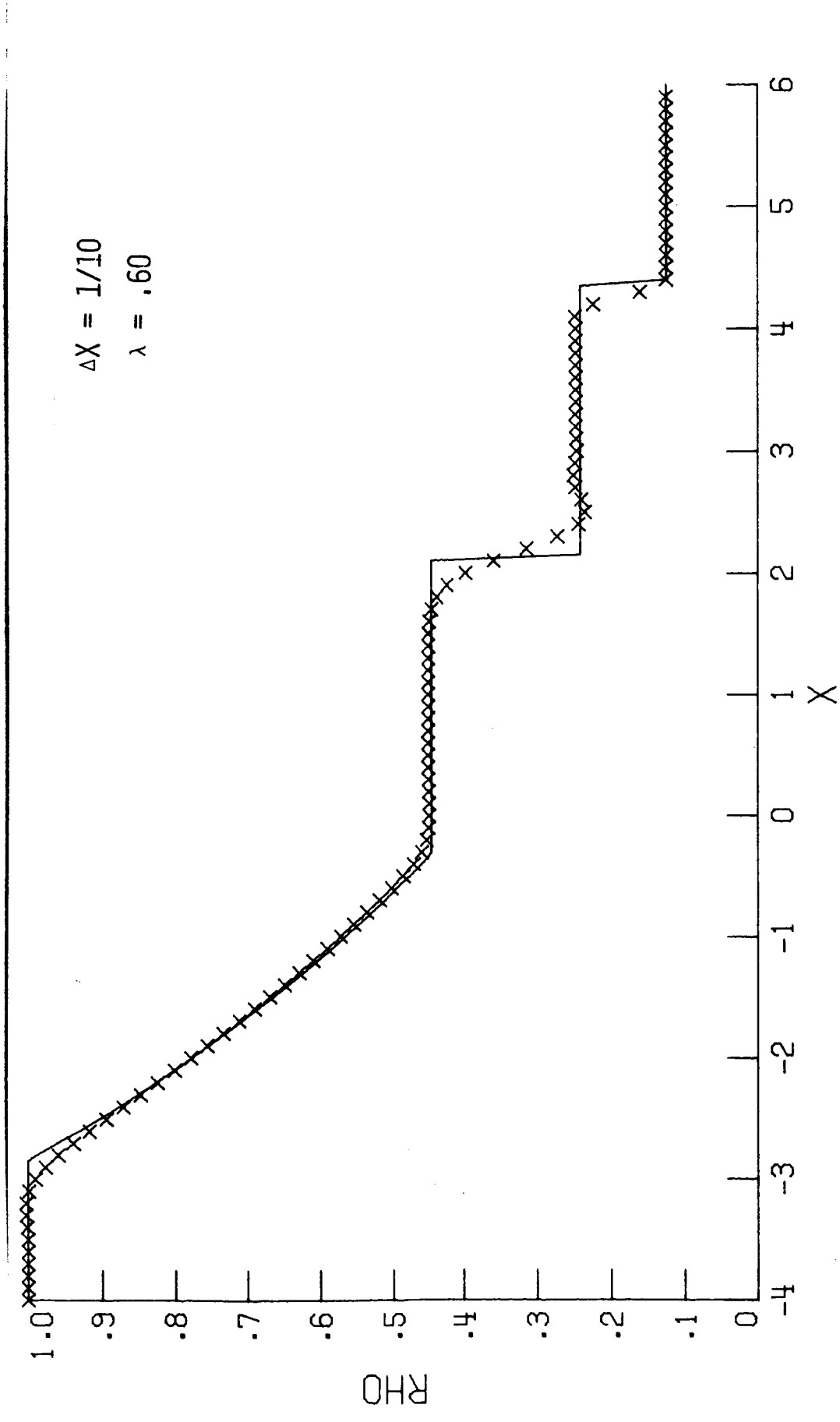

Figure 3a. A comparison of the numerical and exact solution of a Riemann problem; the shock speed exhibits a relative error of 3% (Scheme 1.2).

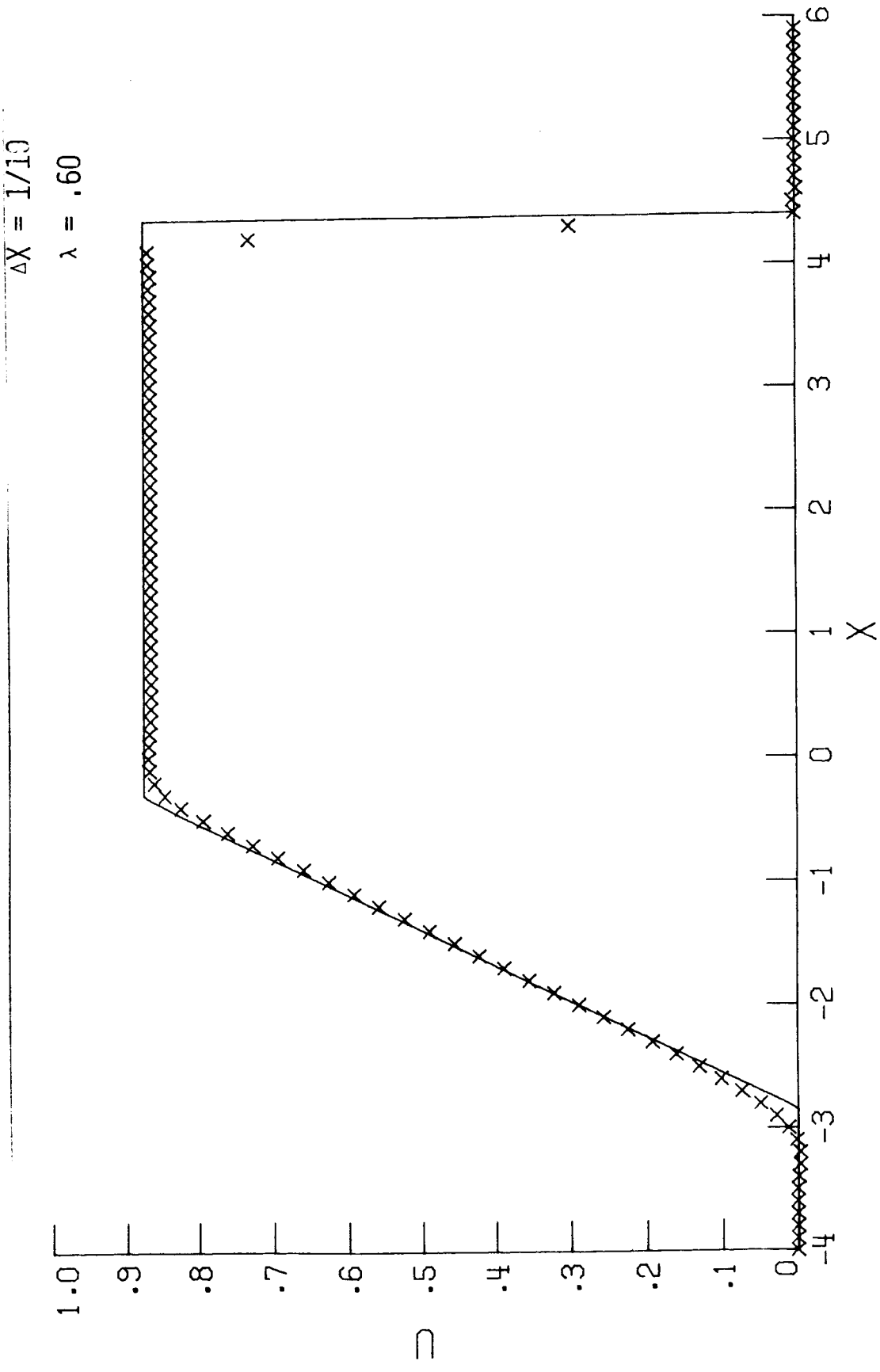

Figure 3b. A comparison of the numerical and exact solution of a Riemann problem; the shock speed exhibits a relative error of 3% (Scheme 1.2).

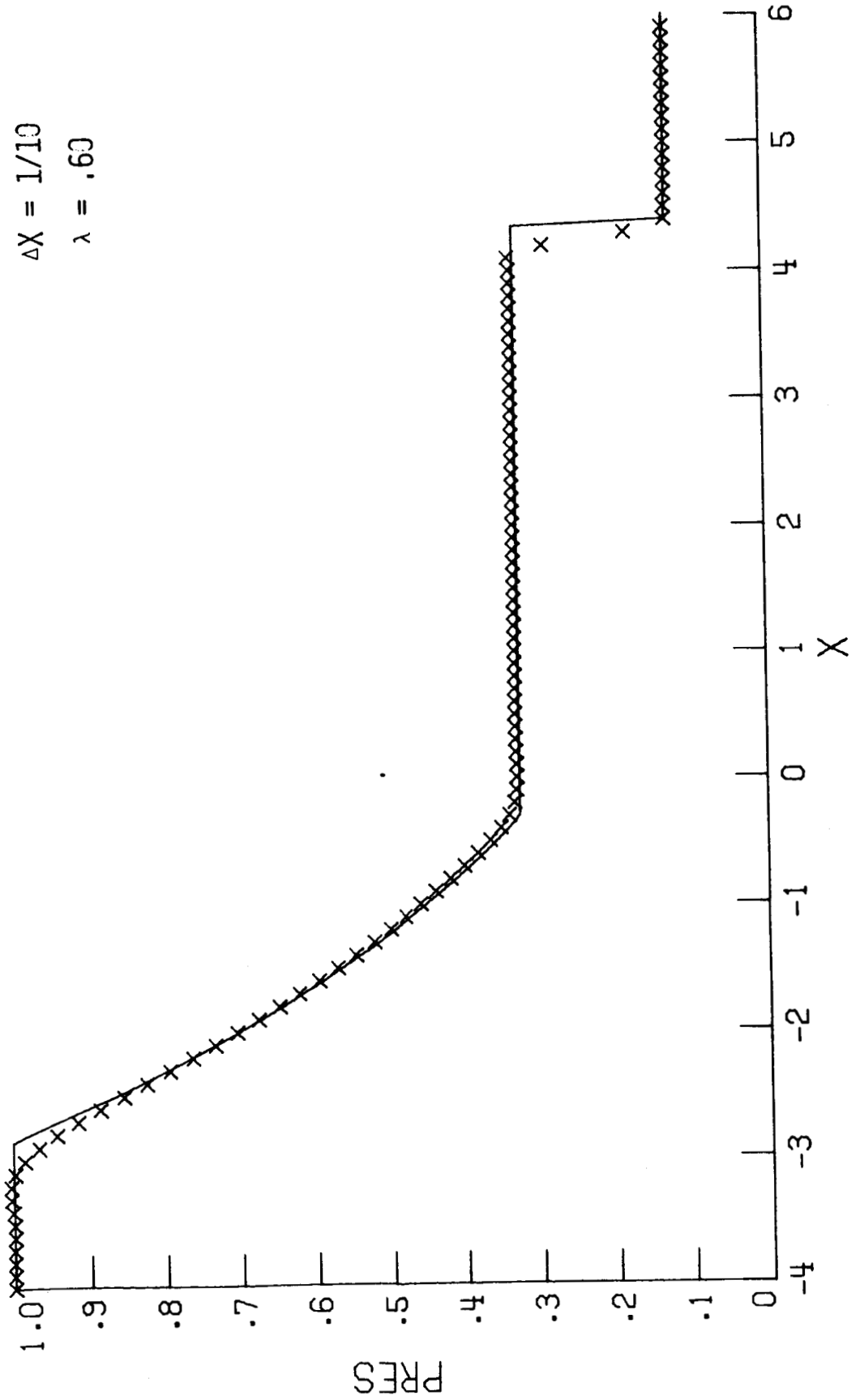

Figure 3c. A comparison of the numerical and exact solution of a Riemann problem; the shock speed exhibits a relative error of 3% (Scheme 1.2).

2. CONVECTIVE-DIFFUSION EQUATIONS

2A. A Compact Scheme

The scalar convective-diffusion equation

$$u_t + au_x + bu_y - \nu\nabla^2 u = 0, \tag{2.1}$$

describes the transport and diffusion of vorticity in two-dimensional, incompressible, viscous flow and serves as a model for the more general Navier-Stokes equations. We shall limit our discussion to the scalar problem partly for expository purposes and partly because the modification required to treat certain types of systems of equations can be readily described in terms of the scalar results.

The dimensionless parameters $\theta_x = a\Delta x/2\nu$, $\theta_y = b\Delta y/2\nu$ play an important role in discussing finite difference schemes for (2.1). Borrowing from fluid dynamics applications of (2.1), $|\theta_x|$ and $|\theta_y|$ are the local cell Reynolds numbers associated with each dimension. The function

$$q(\theta) = \coth\theta - \theta^{-1}, \tag{2.2}$$

will play a fundamental role in our development and for which the following approximations are useful: with $\operatorname{sgn}\theta = \theta/|\theta|$,

$$\begin{aligned} &\text{i)} \quad q(\theta) \simeq \theta/3, && \theta \text{ small} \\ &\text{ii)} \quad q(\theta) \simeq \operatorname{sgn}\theta - \theta^{-1}, && \theta \text{ large.} \end{aligned} \tag{2.3}$$

Finally, we write

$$\hat{q}_x \overset{\text{def}}{=} \frac{\Delta x}{2} q(\theta_x), \quad \hat{q}_y \overset{\text{def}}{=} \frac{\Delta y}{2} q(\theta_y). \tag{2.4}$$

Now write (2.1) as the system

$$\begin{aligned} u_t + au_x + bu_y - (v_x + w_y) &= 0, \\ u_x - v &= 0, \\ u_y - w &= 0. \end{aligned} \tag{2.1}'$$

The origin of the following compact scheme is described in the Appendix:

$$\begin{aligned} &\text{a)} \quad (\delta_t + a\delta_x + b\delta_y)u^n_\bullet - \nu\delta_x v^n_\bullet - \nu\delta_y w^n_\bullet = 0, \\ &\text{b)} \quad \mu_t u^n_\bullet = \mu_x u^n_\bullet = \mu_y u^n_\bullet, \\ &\text{c)} \quad \delta_x u^n_\bullet + \hat{q}_x \delta_x v^n_\bullet = \mu_x v^n_\bullet, \\ &\qquad\quad \delta_y u^n + \hat{q}_y \delta_y w^n_\bullet = \mu_y w^n_\bullet. \end{aligned} \tag{2.5}$$

A proof of the convergence of this scheme again follows by an energy argument which we describe in one dimension. Omitting indices, for $\Delta x \to 0$ we may neglect $\hat{q}_x$ in (2.5c) so that if (2.5a) is multiplied by $\mu_t u$ there results after employing (2.5b,c) $0 = \frac{1}{2}\delta_t u^2 + \nu(\mu_x v)^2 + \delta_x(\frac{a}{2}u^2 - \nu uv)$. Summing on the implied spatial index and assuming homogeneous boundary conditions for u there results $\| u^n\| \leq \| u^0\|$ where $\| u^n\|^2 = \sum (u^n)^2 \Delta x$ and in which the strict inequality holds unless $u = \text{const}$. Thus (2.5) is dissipative. The result implies that $u^n_\bullet$ converges to the solution of (2.1) for any fixed value of the mesh ratio $\lambda_x = \Delta t/\Delta x$.

Arguments similar to that employed in Part 1 for the hyperbolic problem lead to the following:

Two-step Procedure

i) eliminate $u_\bullet^{n+\frac{1}{2}}$ in (2.5a,b) to obtain, with $\tau = \Delta t/2$,

$$
\begin{aligned}
(\mu_x + \tau a\delta_x + \tau b\delta_y)u_\bullet^n - \nu\tau\delta_x v_\bullet^n - \nu\tau\delta_y w_\bullet^n &= u_\bullet^{n-\frac{1}{2}}, \\
(\mu_y + \tau a\delta_x + \tau b\delta_y)u_\bullet^n - \nu\tau\delta_x v_\bullet^n - \nu\tau\delta_y w_\bullet^n &= u_\bullet^{n-\frac{1}{2}},
\end{aligned} \tag{2.6}
$$

which, with (2.5c) leads to an implicit system of equations which, with boundary conditions for $u_\bullet^n$ specified, determines $u_\bullet^n$, $v_\bullet^n$, $w_\bullet^n$.

ii) with the solution $(u_\bullet^n, v_\bullet^n, w_\bullet^n)$ so determined, $u_\bullet^{n+\frac{1}{2}}$ is then obtained from the leapfrog equation (2.5a) or from (2.5b).

2B. Solution Methods

The first of the two-step solution method described above requires the solution of an implicit system of equations. We shall here examine several approaches to this problem.

(i) An ADI Method

The system under discussion is described (2.6) and (2.5c). A more compact description of these equations can be given which shows the close relationship to the hyperbolic case. Let

$$
P_x(\tau) \overset{\text{def}}{\equiv} \begin{pmatrix} \mu_x + \tau a\delta_x & -\nu\tau \\ \delta_x & \hat{q}_x\delta_x - \mu_x \end{pmatrix},
$$

$$
P_y(\tau) \overset{\text{def}}{\equiv} \begin{pmatrix} \mu_y + \tau b\delta_y & -\nu\tau \\ \delta_y & \hat{q}_y\delta_y - \mu_y \end{pmatrix},
$$

and

$$R_x \overset{\text{def}}{\equiv} \begin{pmatrix} a\delta_x & -\nu\delta_x \\ 0 & 0 \end{pmatrix} ,$$

$$R_y \overset{\text{def}}{\equiv} \begin{pmatrix} b\delta_y & -\nu\delta_y \\ 0 & 0 \end{pmatrix} .$$

Then the system (2.6) and (2.5c) is described by

$$\begin{aligned} P_x(\tau)\begin{pmatrix} u_\bullet^n \\ v_\bullet^n \end{pmatrix} + \tau R_y \begin{pmatrix} u_\bullet^n \\ w_\bullet^n \end{pmatrix} &= \begin{pmatrix} u_\bullet^{n-\frac{1}{2}} \\ 0 \end{pmatrix} , \\ P_y(\tau)\begin{pmatrix} u_\bullet^n \\ w_\bullet^n \end{pmatrix} + \tau R_x \begin{pmatrix} u_\bullet^n \\ v_\bullet^n \end{pmatrix} &= \begin{pmatrix} u_\bullet^{n-\frac{1}{2}} \\ 0 \end{pmatrix} . \end{aligned} \tag{2.7}$$

As in the hyperbolic case the solution, to terms of second order in τ, can be obtained from

$$\begin{aligned} \begin{pmatrix} u_\bullet^n \\ v_\bullet^n \end{pmatrix} &= P_x^{-1}(\tau)\left(I - \tau R_y P_y^{-1}(\tau)\right)\begin{pmatrix} u_\bullet^{n-\frac{1}{2}} \\ 0 \end{pmatrix} , \\ \begin{pmatrix} u_\bullet^n \\ w_\bullet^n \end{pmatrix} &= P_y^{-1}(\tau)\left(I - \tau R_x P_x^{-1}(\tau)\right)\begin{pmatrix} u_\bullet^{n-\frac{1}{2}} \\ 0 \end{pmatrix} , \end{aligned} \tag{2.8}$$

in which the solutions represented by P_x^{-1}, P_y^{-1} can be obtained by Keller's method for algebraic two-point boundary value problems.

A discussion of the truncation error for (2.5) follows arguments similar to those for the hyperbolic problem. The elimination of $u_\bullet^{n+\frac{1}{2}}$ in (2.5) led to (2.7); by eliminating $u_\bullet^{n-\frac{1}{2}}$ instead there results

$$P_x(-\tau)\begin{pmatrix} u_\bullet^n \\ v_\bullet^n \end{pmatrix} - \tau R_y \begin{pmatrix} u_\bullet^n \\ w_\bullet^n \end{pmatrix} = \begin{pmatrix} u_\bullet^{n+\frac{1}{2}} \\ 0 \end{pmatrix} ,$$

$$P_y(-\tau)\begin{pmatrix} u_\bullet^n \\ w_\bullet^n \end{pmatrix} - \tau R_x \begin{pmatrix} u_\bullet^n \\ v_\bullet^n \end{pmatrix} = \begin{pmatrix} u_\bullet^{n+\frac{1}{2}} \\ 0 \end{pmatrix} .$$

It then follows that

$$P_x(\tau)\begin{pmatrix} u_\bullet^{n+1} \\ v_\bullet^{n+1} \end{pmatrix} - P_x(-\tau)\begin{pmatrix} u_\bullet^n \\ v_\bullet^n \end{pmatrix} + \Delta t \mu_t R_y \begin{pmatrix} u_\bullet^{n+\frac{1}{2}} \\ w_\bullet^{n+\frac{1}{2}} \end{pmatrix} = 0 ,$$

$$P_y(\tau)\begin{pmatrix} u_\bullet^{n+1} \\ w_\bullet^{n+1} \end{pmatrix} - P_y(-\tau)\begin{pmatrix} u_\bullet^n \\ w_\bullet^n \end{pmatrix} + \Delta t \mu_t R_x \begin{pmatrix} u_\bullet^{n+\frac{1}{2}} \\ v_\bullet^{n+\frac{1}{2}} \end{pmatrix} = 0 .$$

The truncation error as estimated from these expressions is $0(\Delta t^2)$ and is independent of θ_x, θ_y.

ii) A More Efficient Method

The solution of (2.8) involves the solution of one-dimensional problems of the form

$$P U_\bullet^n = \begin{pmatrix} g_\bullet \\ 0 \end{pmatrix} , \tag{2.9}$$

where $P = (P_x, P_y)$ and, accordingly, $U_\bullet^n = (u_\bullet^n, v_\bullet^n)^T$ or $U_\bullet^n = (u_\bullet^n, w_\bullet^n)^T$.

Keller's method applied to (2.9) exploits the fact that the system with its associated boundary conditions can be written as a block tridiagonal system of equations and solved accordingly. We shall now show that the solution for the scalar component $u_\bullet^n$ of $U_\bullet^n$ can itself be obtained as the solution of a scalar tridiagonal system and from which $v_\bullet^n$ or $w_\bullet^n$ can then easily be obtained.

For the hyperbolic limit $(\nu \to 0)$ this tridiagonal system reduces to a bidiagonal system which describes an upwind or downwind differencing method for $u_\bullet^n$ depending upon $\operatorname{sgn} a$. The solution method thus not only reduces the complexity of the computation of $U_\bullet^n$ but also illuminates an aspect of flux-splitting methods which have been employed for hyperbolic equations (c.f. Steger and Warming [15], van Leer [7]).

The details of this reduction are: the equation

$$P_x \begin{pmatrix} u_i^n \\ v_i^n \end{pmatrix} = \begin{pmatrix} g_i^n \\ 0 \end{pmatrix} \quad ,$$

involves the values $u_{i\pm\frac{1}{2}}^n$, $v_{i\pm\frac{1}{2}}^n$ in a cell. Solving for $v_\bullet^n$ in terms of $u_\bullet^n$ yields

$$\begin{aligned} \nu\lambda_x v_{i+\frac{1}{2}}^n &= \tfrac{1}{2}[(1+q_x)(1+\lambda_x a)+\kappa_x]u_{i+\frac{1}{2}}^n + \tfrac{1}{2}[(1+q_x)(1-\lambda_x a)-\kappa_x]u_{i-\frac{1}{2}}^n - (1+q_x)g_i^n, \\ \nu\lambda_x v_{i-\frac{1}{2}}^n &= \tfrac{1}{2}[(q_x-1)(1+\lambda_x a)+\kappa_x]u_{i+\frac{1}{2}}^n + \tfrac{1}{2}[(q_x-1)(1-\lambda_x a)-\kappa_x]u_{i-\frac{1}{2}}^n + (1-q_x)g_i^n, \end{aligned}$$

(2.10)

where $q_x = q(\theta_x)$, $\lambda_x = \Delta t/\Delta x$, and

$$\kappa_x \overset{\text{def}}{=} \frac{2\nu\Delta t}{(\Delta x)^2} = \theta_x^{-1} \cdot a\lambda_x .$$

Setting $k = i+\frac{1}{2}$ and equating the representations for $v_\bullet^n$ given above there results

$$\begin{aligned} (2.11)\ \tfrac{1}{2}[(1-q_x)(1+\lambda_x a)-\kappa_x]u_{k+1}^n + \tfrac{1}{2}[(1+q_x)(1-\lambda_x a)-\kappa_x]u_{k-1}^n \\ + (1+a\lambda_x q_x+\kappa_x)u_k^n = (1-q_x)u_{k+\frac{1}{2}}^{n-\frac{1}{2}} + (1+q_x)u_{k-\frac{1}{2}}^{n-\frac{1}{2}}. \end{aligned}$$

This tridiagonal system may be efficiently solved for $u_\bullet^n$ when boundary values for $u_\bullet^n$ are prescribed. For $|\theta_x| \to \infty$ $\kappa_x \to 0$ and $q \to \operatorname{sgn} a$ so that (2.11) results in

$$\tfrac{1}{2}(1+\lambda_x a)u_k^n + \tfrac{1}{2}(1-\lambda_x a)u_{k-1}^n = u_{k-\frac{1}{2}}^n. \tag{2.12}$$

This may be solved by forward marching when $a > 0$ and by backward marching when $a < 0$. Equation (2.12) exhibits a familiar result for singular perturbation problems: in this case, for $\nu \to 0$, one boundary condition is ineffective for the limiting problem.

The preceding discussion may be extended to systems of equations if a can be written $a = T^{-1}\alpha T$ where T is nonsingular and α is a real diagonal matrix. In this case

$$q(\theta_x) = T^{-1}q(\alpha\Delta x/2\nu)T,$$

and explicit upwind-downwind forms similar to (2.12) again result if characteristic variables are employed.

2C. A More General Problem

Because of its formal similarity to the Navier-Stokes equation it is worth examining the modifications required to treat the more general equation

$$u_t + au_x + bu_y - (cu_{xx} + 2du_{xy} + eu_{yy}) = 0, \tag{2.13}$$

where $ce \geq d^2$.

We now define

$$\theta_x = c^{-1}a\Delta x/2, \qquad \theta_y = e^{-1}b\Delta y/2, \tag{2.14}$$

and

$$
\begin{aligned}
r(\theta) &= \theta^{-1} - (\sinh\theta)^{-1} \\
&\simeq \theta/6, \qquad \theta \text{ small} \\
&\simeq \theta^{-1} \qquad \theta \text{ large.}
\end{aligned}
\tag{2.15}
$$

The approximation (described in the Appendix) which led to (2.5) in the case of eq. (2.1) now leads to

$$
\begin{aligned}
&\text{a)} \quad (\delta_t + a\delta_k + b\delta_y)u_\bullet^n - (\delta_x v^n + \delta_y w_\bullet^n) = 0, \\
&\text{b)} \quad \mu_t u_\bullet^n = \mu_x u_\bullet^n = \mu_y u_\bullet^n , \\
&\qquad (c\delta_x + d\delta_y)u_\bullet^n = (\mu_x - \hat{q}_x\delta_x)v_\bullet^n + \hat{r}(\theta_y)\delta_y w_\bullet^n , \\
&\text{c)} \\
&\qquad (d\delta_x + e\delta_y)u_\bullet^n = (\mu_y - \hat{q}_y\delta_y)w_\bullet^n + \hat{r}(\theta_y)\delta_y v_\bullet^n ,
\end{aligned}
\tag{2.16}
$$

in which we have set $\hat{q}(\theta_x) = \frac{\Delta x}{2}\, c\, q(\theta_x) c^{-1}$, $\hat{q}(\theta_y) = \frac{\Delta y}{2}\, e\, q(\theta_y) e^{-1}$, $\hat{r}(\theta_x) = \frac{\Delta x}{2}\, d\, r(\theta_x) c^{-1}$, and $\hat{r}(\theta_y) = \frac{\Delta y}{2}\, d\, r(\theta_y) e^{-1}$.

The two-step scheme described for (2.5) again applies for (2.16) if the operators R_x and R_y are suitably modified. In particular, P_x and P_y are unchanged so that the solution $u_\bullet^n$ is again determined by an equation of the form (2.11). Thus the occurrence of the mixed derivative term u_{xy} in (2.13) simply leads to added coupling terms in (2.16c).

The fact that the coefficient matrices c, d, e which occur in the Navier-Stokes equations are singular prevents a direct extention of these results to this important problem. The present discussion suggests the type of result which may be expected, however.

2D. Numerical Experiments

Experiment 2.1

The solution $u = \cos(x-t)\exp(-\nu t)$ of $u_t + u_x - \nu u_{xx} = 0$ was computed at $t = 4\pi$ for $\nu = 10^{-2}$ and the values $\Delta x = \pi/20, \pi/40, \pi/80$. The L_1 norm of the numerical errors as a function of $\lambda = \Delta t/\Delta x$ are given in the Table IV.

Table IV

Δx	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
$\pi/20$	1.7×10^{-3}	0.5×10^{-3}	9.3×10^{-3}
$\pi/40$	0.3×10^{-3}	0.27×10^{-3}	2.5×10^{-3}
$\pi/80$	0.8×10^{-4}	0.07×10^{-3}	0.6×10^{-3}

The results confirm the assertion made earlier that (2.5) is second order accurate. Similar results were obtained using the L_2 norm.

Figure 4 indicates steady-state boundary layer profiles ($t = 10$) for $u_t + u_x - \nu u_{xx} = 0$ with boundary conditions $u(0) = 1$, $u(1) = 0$ and initial condition $u(x,0) = (1-x)$; (2.5) was employed with $\Delta x = 1/20$ and $\lambda = 1$. The exact solution is indicated by the solid curves. The numerical results indicate fairly close agreement with the exact solution in the neighborhood of $x = 1$ for various values of ν, the agreement being within the range predicted by the theory both inside and outside of the boundary layer.

Experiment 2.2 - Burger's Equation

Figure 5 describes results for Burger's equation $u_t + (u^2/2)_x = \nu u_{xx}$ with $u(x,0) = 1$ for $x < 0.5$ and $u(x,0) = 0$ for $x > 0.5$ for $\nu = 10^{-2}, 10^{-3}$ at time $t = 1.0$ with $\Delta x = 1/50$ and $\lambda = 1.0$; the vertical line indicates the position of the shock for the limiting value $\nu = 0$.

Figure 6 illustrates the solution of Burger's equation at time $t = 2.0$ (steady-state) with boundary condition $u(0) = 1$, $u(1) = -1$ and initial condition $u(x,0) = 1 - 2x$. The maximum value of the local cell Reynolds number was $R_c = 112.23$.

Experiment 2.3

The two step solution method (2.8) and (2.5a) was employed to calculate the steady-state solution of $u_t + 2u_x + u_y - \nu\nabla^2 u = 0$ which is determined in the unit square by the boundary conditions $u(x,0,t) = 0$, $u(y,o,t) = 10^2$, $u(x,1,t) = 10^2$, $u(1,y,t) = 0$. For $\nu = 0$, the solution is $u = 0$ for $y < x/2$, $u = 10^2$ for $y > x/2$.

The data presented in Figure 7a illustrates solution values obtained when $\Delta x = \Delta y = 0.1$ for $\nu = 10^{-2}$ and 10^{-3}. Figure 7b describes results when $\Delta x = 0.1$, $\Delta y = 0.05$ for $\nu = 10^{-2}$ and $\nu = 10^{-3}$. The oscillations evident behind the limiting ($\nu = 0$) discontinuity in Figure 6 reflect the typical behaviour of a second order method at a discontinuity unless the mesh is adjusted accordingly.

CONCLUDING REMARKS

This paper has described related implicit difference schemes for treating hyperbolic systems of equations and the scalar convective diffusion equation both of which share a common approximation rationale as well as a common solution technique. Numerical evidence indicates that both schemes can be employed to treat nonlinear problems. The accuracy of approximation to the dissipative hyperbolic problem is proportional to an artificial diffusion parameter ε while the approximation to the convective-diffusion is second order accurate and is independent of the value of the local cell Reynolds number. For both schemes conventional energy estimates are available when the coefficients are constant.

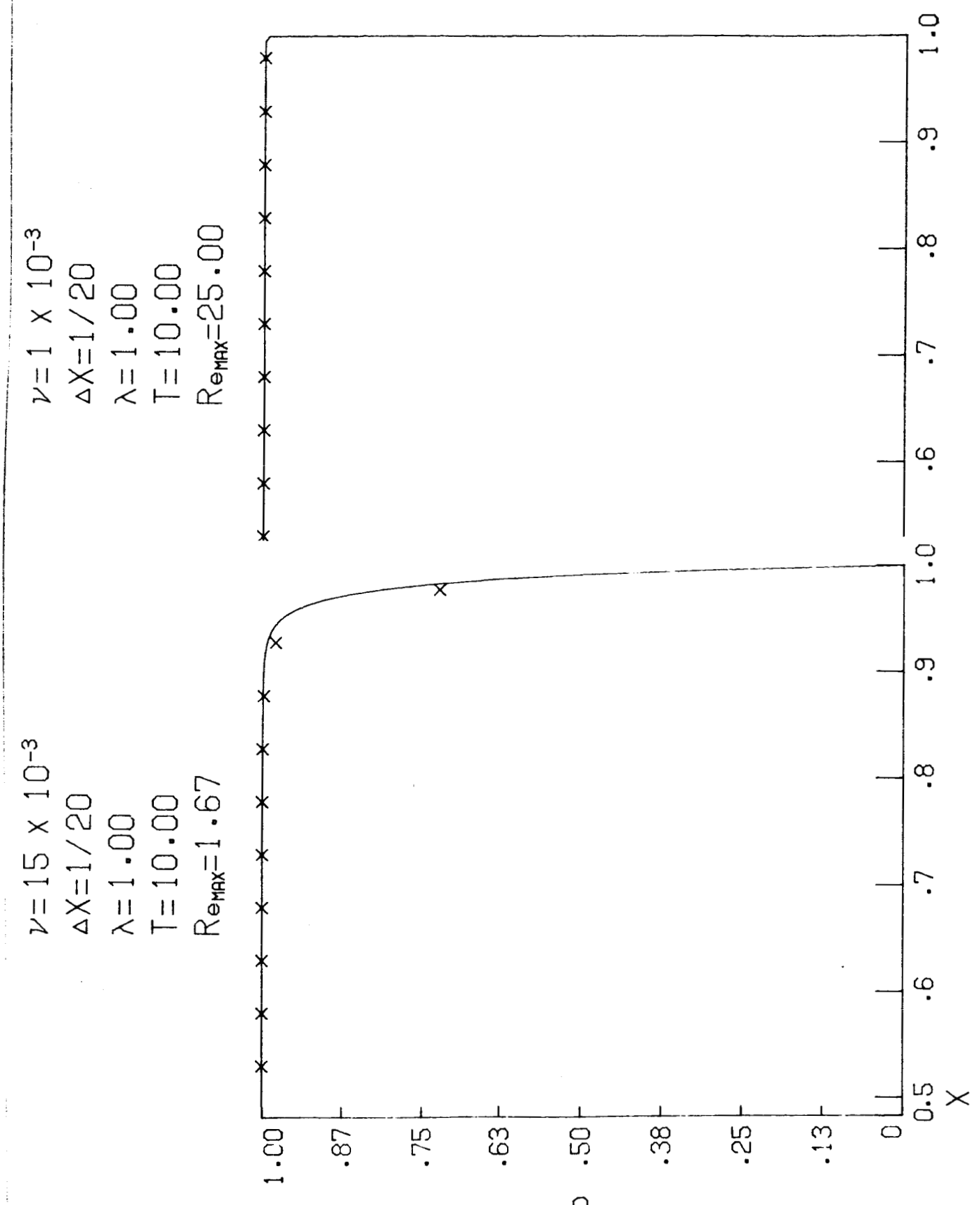

Figure 4. A comparison of boundary layer profiles for various cell Reynolds numbers.

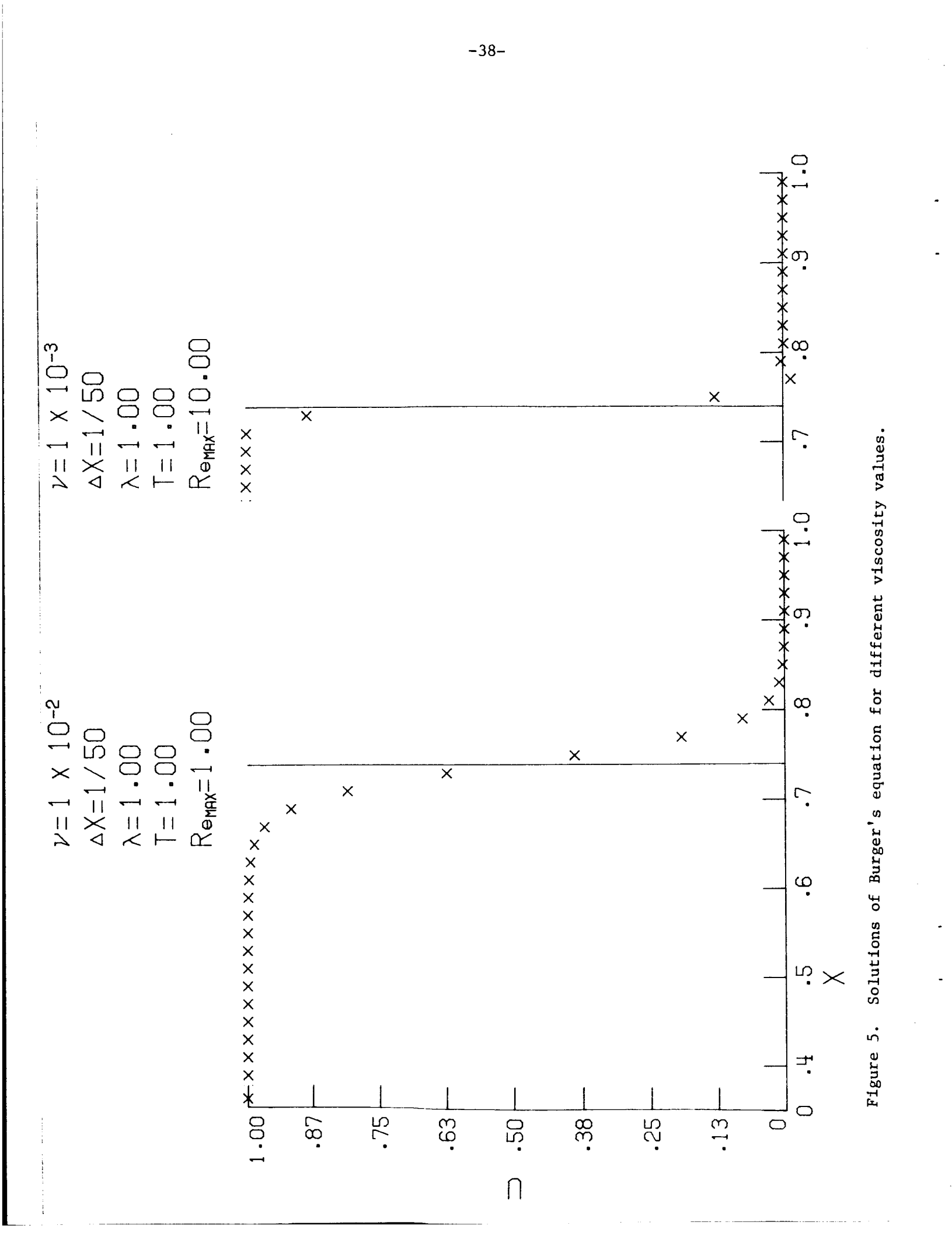

Figure 5. Solutions of Burger's equation for different viscosity values.

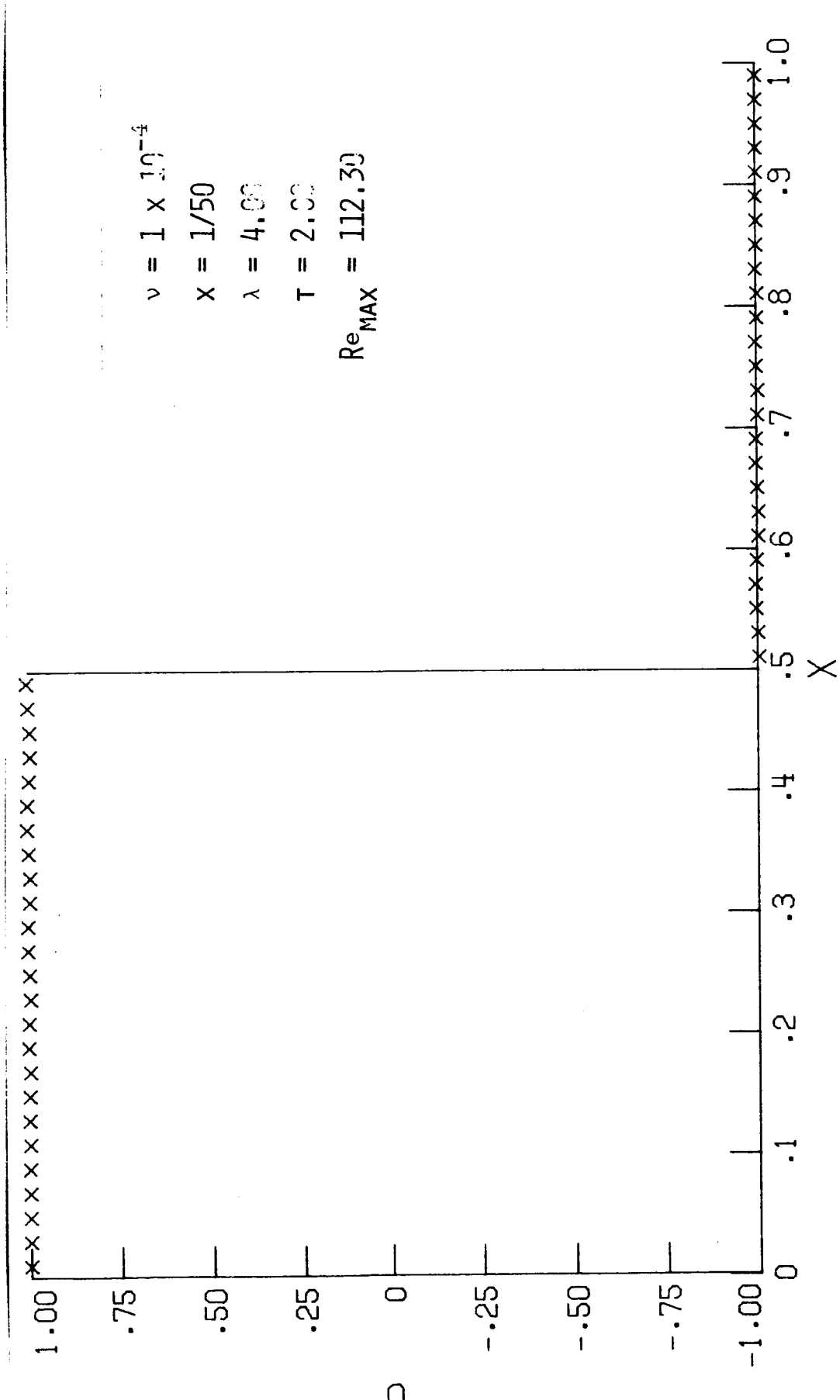

Figure 6. Steady-state solution of Burger's equation.

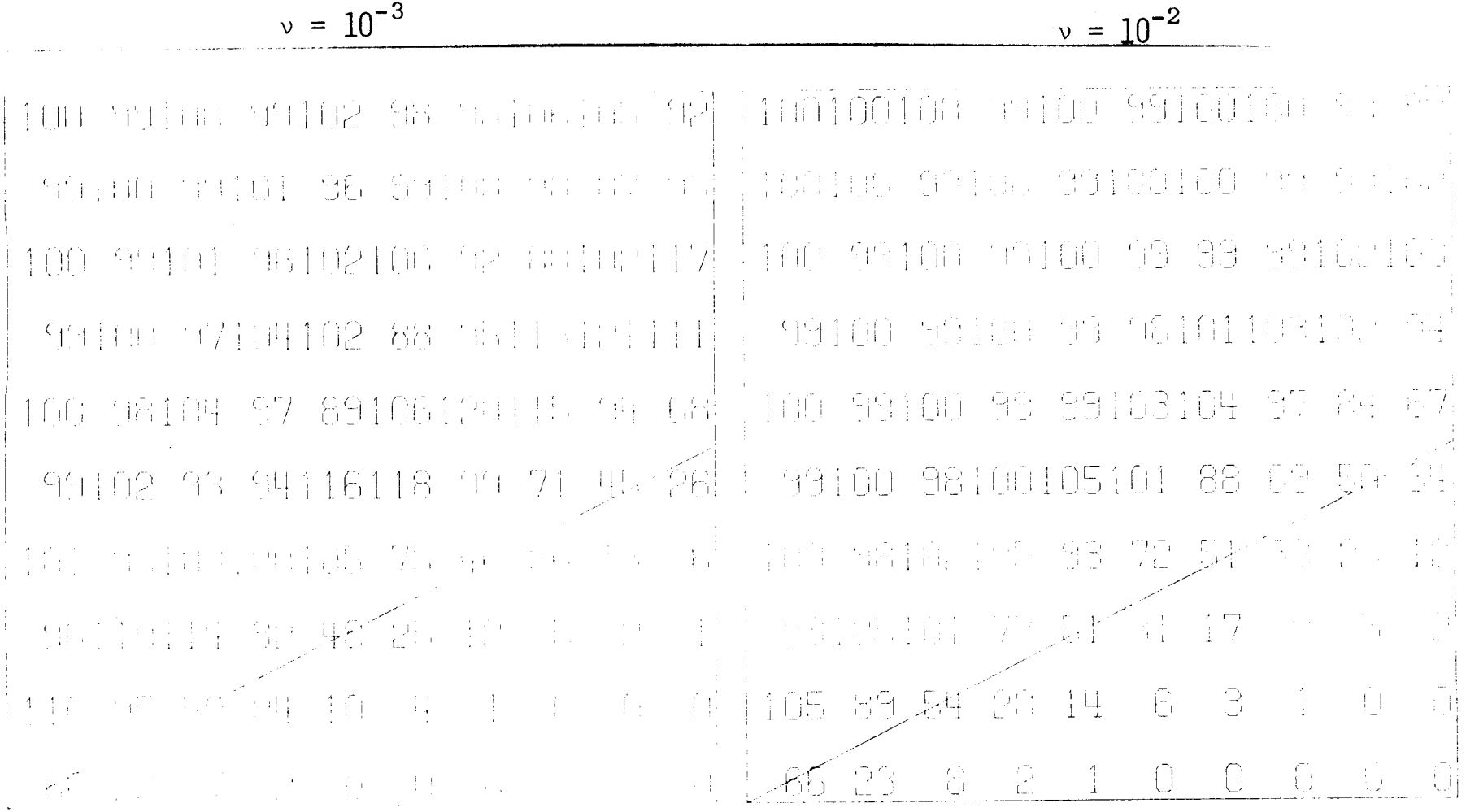

Figure 7a. Steady-state solution of convective-diffusion equation ($\lambda_x = \lambda_y = 1$).

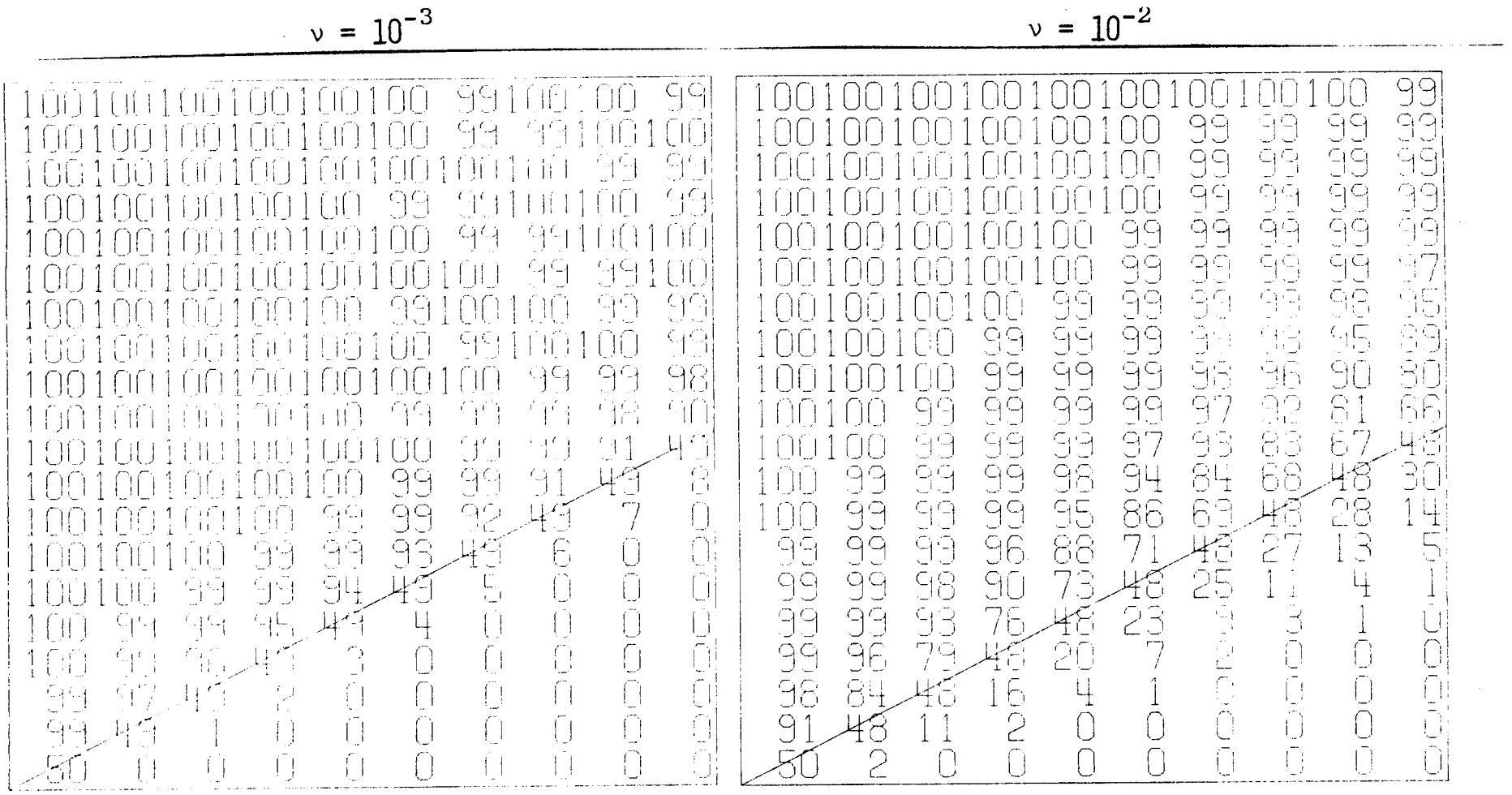

Figure 7b. Steady-state solution of convective-diffusion equation ($\lambda_x = 1, \lambda_y = 2$).

Acknowledgment

We wish to thank J. Flaherty, V. Pereyra, and S. Wornom for assistance about numerical methods underlying the solution of algebraic two-point boundary value problems. Particular thanks are due to J. Oliger for helpful observations about the theoretical development of the paper.

APPENDIX A

The Underlying Approximation Method

The schemes described in this paper have their origin in a common approximation method which we describe for one space dimension.

Divide the fundamental solution domain D $0 < x < 1$, $0 < t < T$ uniformly into $M \cdot N$ rectangular cells each of area $\Delta x \Delta t$. If π_i^n is a typical cell with center point (x_i, t_n) let $\sigma_{i\pm\frac{1}{2}}^n$, $\tau_i^{n\pm\frac{1}{2}}$ denote its vertical and horizontal sides as indicated in Figure 7.

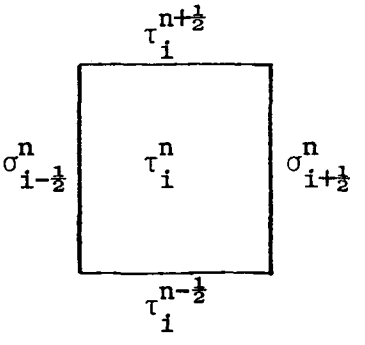

Figure 7

There are thus a total of $(M+1)N + (N+1)M$ sides of which $2(M+N)$ lie on the boundary of D and $2MN - (M+N)$ lie interior to D.

Consider first the equation $U_t + AU_x = 0$. If $U_i^{n\pm\frac{1}{2}}$ and $U_{i\pm\frac{1}{2}}^n$ are the average values of U on the sides $\tau_i^{n\pm\frac{1}{2}}$, $\sigma_{i\pm\frac{1}{2}}^n$ of π_i^n then the MN conditions

$$\delta_t U_i^n + A\delta_x U_i^n = 0, \tag{A.1}$$

will imply that Gauss' theorem holds on the union of any contiguous set of cells. Suppose now that the global solution is approximated by functions which are solutions of the linear differential equation in each cell each of which depends, say, upon α parameters, i.e., if $\phi_1, \phi_2, \cdots \phi_\alpha$ are linearly independent solutions in a cell, we let

(A.2) $$U \simeq \sum_{i=1}^{\alpha} \underline{\phi}_i c_i.$$

If $\alpha = 2$ the mixed initial-boundary value problem (1) may be approximated as follows (c.f. Rose [14]): set $\underline{\phi}_1 = I$, $\underline{\phi}_2 = (xI - tA)$; then the parameters $\underline{c}_1$, $\underline{c}_2$ will be determined by any two of the four average values $U_i^{n\pm\frac{1}{2}}$ and $U_{i\pm\frac{1}{2}}^n$ associated with the sides of π_i^n. Elimination of the parameters yields two relationships between these values one of which is expressed by (A.1) and the other by

(A.3) $$\mu_t U_i^n = \mu_x U_i^n.$$

There thus result $2MN$ conditions for the $2MN + (M+N)$ average values. By imposing the boundary and initial conditions in (1) a determined system of equations results. As we have shown through the use of an energy argument, when A is symmetric and constant this approximation method converges in an L_2 norm for smooth solutions.

If, more generally, we consider the conservation equation $U_t + F_x(u) = 0$, Gauss' theorem requires that

(A.4) $$\delta_t U_i^n + \delta_x F_i^n = 0,$$

where now F_i^n are average flux values on the sides of $\sigma_{i\pm\frac{1}{2}}^n$ of π_i^n. Let $A = \text{grad}\, F$ and consider the following problem: construct a solution $U(x,t:\pi_i^n)$ in each all cell depending upon two parameters such that $U(x,t:\pi_i^n)$ yields the same average values of the flux values $F_{i\pm\frac{1}{2}}^n$ and the same average values $U_i^{n\pm\frac{1}{2}}$ as the solution itself. If

$$U(x,t:\pi_i^n) = \underline{c}_1 + (x - A_i^n t)\underline{c}_2,$$

then $A_i^n \bar{U}_x(\sigma:\pi_i^n) = \underline{c}_2 = F_i^n$ where $\bar{U}_x(\sigma:\pi_i^n)$ indicates the average values

calculated on the side σ. The result of expressing $\underline{c}_1$ and $\underline{c}_2$ in terms of $U_i^{n\pm\frac{1}{2}}$, $F_{i\pm\frac{1}{2}}^n$ is then

$$A_i^n \mu_t U_i^n = \mu_x F_i^n. \tag{A.5}$$

Equations (A.4) and (A.5) obviously state (A.1) and (A.2) in a more general form. In numerical experiments conducted in this study we found no advantage in using (A.4) and (A.5) instead of (A.1) together with the dissipative modification of (A.2) described in (1.2). The numerical results presented in Part I indicate that the dissipative scheme based upon (1.2) appears to provide accurate approximations to nonlinear conservation laws.

For the scalar equation $u_t + au_x - \nu u_{xx} = 0$ similar arguments show that, if $v = u_x$, then

$$\delta_t u_i^n + a\delta_x u_i^n - \nu\delta_x v_i^n = 0, \tag{A.4}$$

expresses Gauss' theorem in terms of the boundary data of cells. Employing (A.2) with $\alpha = 3$ and the elementary solutions

$$\begin{aligned} \underline{\Phi}_1 &= 1 \quad , \\ \underline{\Phi}_2 &= (x - at), \\ \underline{\Phi}_3 &= e^{ax/\nu} \quad , \end{aligned} \tag{A.5}$$

an elimination of parameters yields the scheme (2.5) (1-dimension) which is a determined algebraic system under the given initial-boundary values. Again, the energy argument given earlier establishes the convergence of this approximation scheme when a is constant.

Both schemes employ an approximation basis which consists of wave solutions of the form

$$\phi(\beta,\gamma) = \exp[\beta(x - \gamma t)],$$

where γ is the wave velocity. In the hyperbolic case $\gamma = A$ and the polynomial basis $(1, xI - tA)$ results by setting $\phi_1 = \phi(0,A)$, $\phi_2 = \phi_\beta(0,A)$.

For the convective-diffusion equation the dispersion relationship $\gamma = a - \nu\beta$ holds. In addition to the polynomial solutions $(1, x - at)$ the function $\exp(ax/\nu)$ provides another linearly independent solution for which $\gamma = 0$ when $\beta = a/\nu$. The basis $(1, \exp(ax/\nu))$ is composed of solutions of the steady-state equation $au_x - \nu u_{xx} = 0$ and the method described above can be used to directly provide a difference scheme for the time-independent problem. The result is described by the two equations (2.5a) and (2.5b) in which the term $\delta_t u_i^n$ is set equal to zero. For the time-independent problem this same basis can be used to construct a Green's function on each overlapping subinterval $[x_{i-1}, x_{i+1}]$ having its singularity at $x = x_i$, $x_{i-1} < x_i < x_{i+1}$, a technique which leads to a positive definite tridiagonal difference scheme for Stürm-Liouville problems as shown in Rose [12]. An equivalent point of view has been independently developed and applied to similar singular perturbation problems which arise from steady-state problems (for example, c.f. Berger, et al. [1], El-Mistikawy and Werle [3], Il'in [8], Rose [13]). In this sense the methods described in this paper appear to provide the appropriate extention of such Green's function techniques to time-dependent problems.

For the convective-diffusion equation a polynomial approximation basis also results by taking

$$\phi_1 = \phi(0,a) = 1$$

$$\phi_2 = \phi_\beta(0,a) = x - at$$

$$\phi_3 = \tfrac{1}{2}\,\phi_{\beta\beta}(0,a) = \tfrac{1}{2}(x - at)^2 + \nu t.$$

The difference scheme which is the consequence may be obtained from (II.3) by setting $q = 0$. In view of earlier remarks this basis can be expected to provide an accurate approximation only when the cell Reynolds number $|\theta|$ is small.

References

[1] A. E. Berger, J. M. Solomon, and M. Ciment, *Uniformly accurate difference methods for a singular perturbation problem*, Proc. Intl. Conference on Boundary and Interior Layers, Computational and Asymptotic Methods, J. J. H. Miller, ed., Boole Press, Dublin, 1980, pp. 14-28.

[2] C. deBoor and R. Weiss, *Solveblok: A package for solving almost block diagonal linear systems, with applications to spline approximation and numerical solution of ordinary differential equations*, MRC Technical Report No. 16-75, University of Wisconsin Press, Madison, WI.

[3] T. M. El-Mistikawy and M. J. Werle, *Numerical method for boundary layer with blowing the exponential box scheme*, AIAA J., Vol. 16, 1978, pp. 749-751.

[4] J. E. Flaherty and W. Mathon, *Collocation with polymial and tension splines for singularly perturbed boundary value problems*, SIAM J. on Sci. and Stat. Computing, Vol. 1, No. 2, June 1980, pp. 260-289.

[5] H. B. Keller, *Numerical methods in boundary layer theory*, Annual Rev. Fluid Mech., Vol. 10, 1978, pp. 417-433.

[6] H. B. Keller, *Numerical solution of two-point boundary value problems*, Regional Conference Series in Applied Mathematics, 24, SIAM, 1976.

[7] B. van Leer, *Upwind differencing for hyperbolic systems of conservation laws*, Numer. Meth. for Eng., Vol. I, Dunod, Paris, 1980.

[8] A. M. Il'in, *Differencing scheme for a differential equation with a small parameter affecting the highest derivative*, Mat. Azmetki, Vol. 6, 1969, pp. 237-248, Math. Notes, Vol. 6, pp. 596-602.

[9] M. Lentini and V. Pereyra, *An adaptive finite difference solver for nonlinear two-point boundary problems with mild boundary layers*, SIAM J. Numer. Anal., Vol. 14, 1977, pp. 91-111.

[10] A. R. Mitchell, *Computational Methods in Partial Differential Equations*, Wiley, London, 1969.

[11] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, Interscience, New York, 1967.

[12] M. E. Rose, *Finite difference schemes for differential equations*, Math. Comput., Vol. 18, No. 86, 1964, pp. 179-195.

[13] M. E. Rose, *Weak element approximation to elliptic differential equations*, Numer. Math., Vol. 24, 1975, pp. 185-204.

[14] M. E. Rose, *A "unified" treatment of the wave equation and the Cauchy-Riemann equations*, SIAM J. Numer. Anal., Vol. 18, No. 2, 1981, pp. 372-376.

[15] J. L. Steger and R. F. Warming, *Flux vector splitting of the inviscid gas dynamic equations with applications to finite difference methods*, NASA Technical Memo 78605, July, 1979.

[16] B. Wendroff, *On centered difference equations for hyperbolic systems*, J. Soc. Indust. Appl. Math., Vol. 8, No. 3, 1960, pp. 549-555.

[17] B. Wendroff, *The structures of certain finite difference schemes*, SIAM Review, Vol. 3, No. 3, 1961, pp. 237-242.